Dark Matter Structures and the Free Streaming Scale

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Aurel Schneider

von

Dörflingen SH

Promotionskomitee

Prof. Dr. Ben Moore (Vorsitz)
Prof. Dr. Jürg Diemand

Zürich, 2012

Aurel Schneider
Institute for Theoretical Physics
University of Zurich
Winterthurerstrasse 190
CH-8057 Zürich
Switzerland
aurel@physik.uzh.ch

Contents

Acknowledgments

The Institute for Theoretical Physics at the University of Zurich is a wonderful place for doing a PhD, and I want to thank everyone contributing to this unique working atmosphere.

I am especially grateful to my supervisor Ben Moore, who gave me the opportunity to work on various different research projects and was always there when needed. His numerous advices have been crucial for accomplishing this thesis. I also want to thank Robert Smith for co-advising the thesis, for sharing his expertise, and for a very pleasant collaboration.

I very much appreciated to collaborate with Doninno Anderhalden, Gianfranco Bertone, Jürg Diemand, Lawrence Krauss, Andrea Macciò, Doug Potter, Darren Reed, and Joachim Stadel, and I am grateful for many inspiring discussions. Special thanks go to Doug Potter for always being patient and for helping with various computational problems.

Finally I want to thank my small new family for simply being here and for sharing the not always easy life of a researcher.

Abstract

This thesis is about various aspects of dark matter structure formation under the influence of the velocity free streaming effect. Depending on the nature of dark matter, the free streaming can act on very different scales, going from the AU scale for a usual cold dark matter particle up to the scale of dwarf galaxies for a warm dark matter candidate. The effect of the free streaming may be observable and could give important hints on the nature of the dark matter particle.

In a first part of the thesis we are looking at the disruption processes acting on the smallest earth-mass haloes in a 100 GeV neutralino dark matter scenario. The existence of such microhaloes in the solar neighborhood would be very interesting for direct and indirect dark matter detection experiments. We however show that tidal stripping as well as the disk passages destroy the major part of the microhalo, and that the effect on direct and indirect detection is negligible.

In a second part we are considering a warm dark matter cosmology, studying nonlinear power spectrum, mass function, bias, and concentrations. We find a surprising behavior of the mass function and the concentration, and we interpret it as a consequence of a reversed hierarchy of structure formation.

In the final part of the thesis we study the cylindrical collapse within tidal streams. In a cylindrical geometry the Jeans length is increased with respect to a homogeneous case, and structure formation is suppressed more efficiently. We find that tidal streams do not gravitationally fragment, since they are not cold enough to allow substantial growth of perturbations.

Zusammenfassung

Diese Thesis handelt von unterschiedlichen Aspekten der kosmischen Strukturformation, welche dominiert wird von dunkler Materie. Zentraler Bestandteil ist dabei der sogenannte "free-streaming" Effekt, der ein Ausdruck für die Geschwindigkeitsdispersion eines kollisionsfreien Fluides ist und zu einer Unterdrükung der Strukturbildung auf kleinen Skalen führt. Wie klein diese kritische free-streaming Skala ist, hängt ab von der Art des hypothetischen dunklen Materie-Teilchens und liegt im Bereich zwischen der Grösse des inneren Sonnensystemes für "kalte" dunkle Materie (CDM) und der Grösse von Zwerggalaxien für "warme" dunkle Materien (WDM). Der free-streaming Effekt, könnte unter Umständen beobachtbar sein und würde dann Aufschluss geben, über die Art des hypothetischen dunkle Materie-Teilchens.

In einem ersten Teil der Thesis betrachten wir den zerstörerischen Einfluss des galaktischen Halos auf die kleinsten Strukturen in einem CDM Szenario mit einem 100 GeV Neutralino. Diese Strukturen werden "Mikrohalos" genannt und sind nur etwa eine Erdmasse schwer. Für direkte und indirekte Detektionsexperimente von dunkler Materie wäre die Existenz von Mikrohalos im oder um das Sonnensystems von grosser Bedeutung. Wir zeigen jedoch, dass Gezeitenkräfte und Gravitationswechselwirkung mit Sternen einen Grossteil der Mikrohalos zerstört, oder so stark komprimiert, dass sie keinen Einfluss mehr auf Detektionsexperimente haben.

In einem zweiten Teil betrachten wir eine WDM Kosmologie, wobei wir die nichtlineare spektrale Leistungsdichte, die Massenfunktion und die Konzentration der Halos, sowie den "bias" untersuchen. Die Massenfunktion und die Konzentration zeigen ein überraschendes Verhalten, welches als Zeichen einer Umkehrung der hierarchischen Strukturbildung interpretiert werden kann.

In dem dritten und letzten Teil der Thesis betrachten wir den Gravitations-Kollaps innerhalb eines Zylinders und wenden dieses Modell an auf Gezeitenströme (tidal streams) von Substrukturen in einem galaktischen Halo. Wir zeigen, dass "tidal streams" auf Grund ihrer grossen Geschwindigkeitsdispersion stabil sind und nicht fragmentieren.

1

Introduction

This thesis is a summary of my research activities of the last three to four years. It is not a self consistent work nor does it contain one single research hypothesis, but it is a patchwork of different research projects I have worked on during the time of my PhD. There are however common points relating the different projects, which are the dark matter structure formation and, more specifically, the velocity free streaming effect.

The thesis consists of a general introduction into the research field followed by three published papers and ending with an outlook to future work. I will now give a short overview of the research done in these papers, and discuss how the different subjects are related to the thesis as a whole.

The **first paper** is about the local dark matter distribution within a ΛCDM cosmology. This is of great interest for the search of the dark matter particle, since the interpretation of detection experiments depends on the local dark matter density. Because of the very nonlinear behavior of small scale structure formation and the finite resolution of N-body simulations, it is however impossible to determine the precise distribution at the local scale of our solar system. We perform analytical approximations and toy model simulations to estimate the local dark matter distribution. We thereby assume a hierarchical structure formation, which means that the smallest structures, called microhaloes, form very early, have high concentrations and are therefore the most susceptible to survive the disrupting forces within the Milky way halo. The mass of the microhaloes is determined by the primordial CDM free streaming scale which depends on the nature and mass of the dark matter particle candidate. Assuming a 100 GeV neutralino the microhalo mass is extremely small, comparable to the one of the earth, i.e. $10^{-6} M_{\odot}$. In order to determine the influence of microhaloes on the dark matter distribution, we estimate their number density in the Milky Way halo as well as their resistance to disruption effects, such as tidal streaming and interactions with stars in the galactic disk. We find that only the very centers of the microhaloes survive and more than 95% of the initial mass is

stripped away. The boosting effect of this remnants on direct and indirect detection stays negligibly small. We also explore the local velocity distribution which could show signatures from the initially very cold tidal streams of disrupted microhaloes. The streams are however heated up by stars in the disk, and spikes in phase space are completely washed out, leading to a quasi Maxwellian velocity distribution.

In the **second paper** we study the structure formation of a ΛWDM cosmology. In this alternative scenario the dark matter particle is much lighter than in ΛCDM, yielding a larger free streaming scale and a stronger small scale suppression. The smallest collapsed structures are not earth-mass microhaloes as in ΛCDM, but haloes of about the size of dwarf galaxies, depending on the specific nature and mass of the WDM particle candidate. We explore the ΛWDM cosmological structure formation, using the analytical halo model combined with numerical simulations, and we find important deviations from the usual picture of hierarchical collapse. First of all the halo mass function is considerably shallower than predicted by the EPS approach. Second the concentration mass function is not monotonically increasing towards small masses, but it turns over at some characteristic mass and decreases again. This can be interpreted as an inversion of hierarchical structure formation, at least if one believes the usual explanation that the concentration is set by the background density at the time of halo virialisation. By correcting the mass function and the concentration-mass behavior with corresponding fitting functions, we improve the halo model of WDM considerably, reducing the error on the power spectrum down to 5%, what is comparable to the CDM halo model.

The **third paper** explores the gravitational stability of tidal dark matter streams. In order to model the evolution of density perturbations within a tidal stream, we use the approximation of an infinitely long, expanding cylindrical matter distribution. In this geometry the linear growth and the Jeans criterion is qualitatively different from the case of a homogeneous fluid, and the evolution of gravitational overdensities has to be worked out independently. The growing mode solution of cylindrical collapse is very sensible to the value of the velocity dispersion as well as the amount of longitudinal expansion. Tidal streams expand linearly on average, and their velocity dispersion can be determined from the kinetics of the adjacent substructure. Using these approximations, we find that dark matter (or stellar) tidal streams cannot re-collapse to build small haloes. Perturbation modes within streams are either oscillating or freezing out after an unsubstantial phase of growth. The initial assumption of a linearly expanding cylindrical matter distribution is however far from being a perfect approximation for a real stream, which lies on a bended orbit, is often flattened and oscillates in width and length during one orbital passage. We analyze this effects as well and conclude that they cannot enforce the growth of perturbations.

The common point of all three thesis subjects is the structure formation and evolution at scales, where the velocity dispersion of the fluid starts to act as a suppression mechanism against gravitational collapse. Either the free streaming effect has erased preexisting density perturbations as in the first two papers, or the structure formation is suppressed due to Jeans criterion as in the third paper. Both, the free streaming and the Jeans scale are tightly related, a fact that is explained in section 2.3.3.

The thesis is structured as follows: In chapter 2 we talk about the theoretical background, starting with a short introduction to the big bang cosmology and a discussion on dark matter. We then introduce the concept of linear and nonlinear structure formation, with a special focus on the connection between free streaming and Jeans length as well as on numerical simulations and the extended Press-Schechter approach. In chapter 3 we talk about microhaloes, the smallest structures of a typical CDM cosmology. The Disruption processes acting on microhaloes in the galactic halo are discussed in detail as well as the consequences of microhaloes on dark matter detection experiments. Chapter 4 introduces the alternative WDM cosmology, and combines the outcome of the analytical halo model with numerical simulations. In chapter 5 we then analyse the stability of tidal streams. Finally, chapter 6 is an outlook on future work.

2

THEORETICAL BACKGROUND

2.1. Cosmic history

The standard model of cosmology is based on the picture of an expanding universe, consisting of a hot primordial plasma, which cools down and starts to cluster, building structures on various scales, including our galaxy. The underlying dynamics are governed by Einstein's field equations

$$G_{\mu\nu}(g_{\mu\nu}) = 8\pi G T_{\mu\nu} + g_{\mu\nu}\Lambda, \qquad G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R, \tag{2.1}$$

relating the geometry of space-time (described by the the Ricci tensor $R_{\mu\nu}$, the Ricci scalar R, and ultimately the metric $g_{\mu\nu}$) to its energy content (given by the stress energy tensor $T_{\mu\nu}$ and the cosmological constant Λ). At very early times and on large scales the matter distribution of the universe is assumed to be nearly homogeneous and isotropic. Under this assumption, a class of solutions to the field equations is the Friedmann-Lemaître-Robertson-Walker metric with the line element

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a(t)\left[(1-kr^2)^{-1}dr^2 + r^2 d\theta^2 + r^2\sin\theta d\phi^2\right], \tag{2.2}$$

where k is the curvature parameter and a is a function of cosmic time, known as the scale factor. The behavior of a can be determined by implementing (2.2) into (2.1), yielding the Friedmann equations for an expanding universe:

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3}, \tag{2.3}$$

$$2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = -8\pi G p + \Lambda. \tag{2.4}$$

Here we have assumed a stress energy tensor of the form $T_{\mu\nu} = (\rho + p)u_\mu u_\nu + p g_{\mu\nu}$, where u_μ is the four-vector for the velocities. The Friedmann equations consist of two

coupled, nonlinear differential equations of the scale factor, describing the expansion history of the homogeneous and isotropic universe.

The matter density ρ and pressure p are related by the equation of state, which is usually parametrised as $p = w\rho$. Recent measurements [1, 2] point towards a flat universe, where the curvature parameter k is zero, and we will neglect curvature in what follows.

The Friedmann equations can be used to derive a relation for the energy conservation

$$\dot{\rho} = \frac{\dot{a}}{a}(3p + \rho). \tag{2.5}$$

In the case of a constant equation-of state parameter w, the solution to Eq. (2.5) is $\rho = \rho_0 a^{-3(1+w)}$. For radiation (and highly relativistic matter) $w = 1/3$ and the energy density behaves as $\rho_{\rm hr} \propto a^{-4}$. Non relativistic matter on the other hand has negligible pressure ($w = 0$), and thus $\rho_{\rm nr} \propto a^{-3}$. An expanding universe must therefore have been in an early epoch of radiation domination, followed by an epoch dominated by non relativistic matter. At some point matter domination is again replaced by what we call the epoch of dark energy, since the decreasing density term is overruled by the cosmological constant (see Eq. 2.3).

It is now straight forward to solve the Friedmann equations for the three main epochs in cosmic history. During radiation domination we get $a(t) \propto t^{1/2}$ and during matter domination $a(t) \propto t^{2/3}$. At the epoch of dark energy $a(t)$ asymptotically approaches $a(t) \propto \exp(\sqrt{\Lambda/3}t)$ as the matter density goes to zero. The expansion is therefore decelerating during the epochs of radiation and matter domination, and it changes to an accelerated phase at the late epoch of dark energy domination.

In the very early universe during the epoch of radiation domination, all matter components are in a very hot and dense state, interacting with each other and building a thermodynamical equilibrium. Expansion of space makes this primordial 'soup' of particles cool down, what leads to more and more phase transitions of high energy interactions. An incomplete but chronological list is the transition from free quarks to baryons, the electron-positron annihilation, the neutrino decoupling, the nucleosynthesis of the first atoms, and finally the recombination, where the free electrons are trapped by the nuclei to form the first neutral atoms. With the binding of free electrons, the universe becomes transparent to photons, which freely propagate until today and build what is known as the 'cosmic microwave background' (CMB). Observing the CMB is the furthest look we can make in distance and time, and it is a fundamental constraint on the observable size of the universe.

At about the time of recombination the universe enters the epoch of matter domination and gravitational structure formation starts to become important. This is the beginning of what we call the 'dark ages', where the primordial overdensities first grow linearly, then collapse in a complicated nonlinear way to form the cosmic web and the first virialised objects. The formation of the first stars denotes the end of the dark ages and the beginning of the reionisation period, where the new star light ionises the neutral hydrogen of the intergalactic medium.

In the meantime the complex process of nonlinear structure formation continues in a bottom up fashion, where the first small dark matter haloes merge to build a large variety of structures of different scales. Within this clustering process of dark

matter, the baryons interact electromagnetically, loose energy through radiation, and sink to the bottom of the gravitational potentials, where the densities become high enough to permit a large amount of star formation. Galaxies of various sizes start to populate the universe and trace the underlying density field of the dominant dark matter.

Today the universe has entered the epoch of dark energy, where the main energy component comes from the cosmological constant and the expansion is accelerated. In the future all unbound objects will be redshifted away from each other and we will eventually end up having single galaxies as island universes surrounded by empty space [3].

2.2. Dark Matter

The first hint for dark matter goes back to the early thirties of the last century when Fritz Zwicky studied the velocities of galaxies within the Coma cluster and, due to kinematical reasons, postulated a dominant amount of invisible matter. This picture got widely accepted only about forty years later, when the study of milky way satellites [4, 5] as well as measurements of rotation velocities of galaxies [6] showed the necessity of very massive haloes of dark matter surrounding the visible galaxies. Candidates for the dark matter went from MACHO's (massive astrophysical compact halo objects made of baryons, for example brown dwarfs or black holes, [7]) to elementary particles such as massive neutrinos [8]. Soon it became clear that baryons could not be the dark matter, since the expected abundances of elements from nucleosynthesis give a strong constraint on the baryon density in the universe, which is well below the density needed for a flat or nearly flat universe, predicted by observations. Neutrinos on the other hand emerged as a natural dark matter candidate, since experiments suggested them to have non zero mass [9]. However, after working out the neutrino structure formation, it became clear that the collapse time of neutrino dominated galaxies is to high and that a much colder (or heavier) dark matter particle is needed [10].

Together with the various theories extending the standard model of particle physics there are also several possible dark matter candidates that fulfill the necessary conditions of being neutral, heavy, and interacting only very weakly with other particles. They got named WIMPs, standing for weakly interacting massive particles.

The major drawback of all possible dark matter candidates is that they have not been detected yet, at least not independently of their very visible effect on gravity. Today, a lot of experimental efforts go into the detection of WIMP candidates, where the usual strategy is to measure collisions with baryons, which is a very difficult task, due to the extremely small weak interaction cross section.

Before going a bit more into the details of different dark matter particle candidates, we will summarise the major astrophysical reasons for postulating dark matter. There are numbers of observational evidence for dark matter on a wide range of scales from individual galaxies to the CMB. Some of the most important ones are listed in the following [for more details see 11, 12, and references therein] :

Galaxy rotation curves: The rotation of edge-on galactic disks can be measured up to good accuracy with spectral methods. The observations of a large number of galaxies all indicate flat rotation curves over a large range in radius. This can only be explained by either changing the laws of gravity or by postulating a massive and extended halo of dark matter.

Galaxy clusters: As Zwicky pointed out in the early thirties, the velocities of orbiting galaxies in clusters are too high to be explained by the visible cluster mass. Moreover, strong lensing data indicate an excess of mass in agreement with an extended dark matter halo surrounding galaxy clusters. A very strong observational hint for dark matter comes from the Bullet cluster, which actually consists of two colliding galaxy clusters. Independent measurements of the amount of gas (via X-ray measurements) and the amount of total cluster mass (via lensing) indicate a spatial separation of gas and dark matter, coming from the fact that gas is slowed down through pressure forces, which on the other hand do not affect dark matter. The Bullet cluster is widely considered as the most direct indication for the existence of non baryonic dark matter and disfavors alternative gravity theories such as MOND [13].

CMB constraints: The CMB temperature anisotropies show the imprint of baryonic acoustic oscillations (BAO), which are a direct consequence of gas pressure counteracting the growth of baryon perturbations. Measuring the position and the width of the BAO peaks clearly indicates the existence of a dominating dark matter component.

In summary, there is very strong evidence for dark matter on a variety of scales, and although no dark matter particle has been directly detected yet, it turns out to be extremely hard to explain the above listed phenomena without any form of dark matter.

We will now take a closer look on the nature of the different WIMP dark matter candidates. As already mentioned, a particle candidate must be neutral and heavy with a very small weak interaction cross section. No such candidate appears within the standard model of particle physics, but several theories covering physics beyond the standard model naturally contain a dark matter particle.

2.2.1. Cold dark matter (CDM)

The most popular candidate of dark matter, and the key example for a WIMP, is the lightest supersymmetric particle, called neutralino. It has a mass of about 10 GeV to 1 TeV and only interacts via the weak forces, what makes it a perfect dark matter candidate. A great ongoing effort is devoted to the very difficult detection of the neutralino through nucleon scattering events. A popular detection setup consists of some large and well shielded target crystal, where the interaction with a neutralino produces phonons in the crystal matrix [14, 15]. Other setups are based on shielded tanks of liquid rare metals, where neutralino interactions should produce small charge and light signals [16]. With the existing experiments no WIMP detection has been made up to now, but future experiments with larger targets are on the way, and they will explore much more of the possible parameter space.

Since the neutralino is a Majorana fermion, it can self annihilate and produce a cascade of gamma ray signals. A confirmation of such a signal would be an indirect dark matter detection and could give insight into the dark matter distribution on the sky. Since the strength of the signal depends on the self annihilation cross section as well as on the dark matter density squared, the strongest signal is expected at potential high density regions like the center of our galaxy. Unfortunately, there are a lot of astrophysical gamma ray sources (supernovae, neutron stars, black holes, etc), which are also located at high density regions, and it is very difficult to disentangle between a potential signal and the background [17].

Another often mentioned CDM candidate is the axion, a hypothetical particle postulated to solve the strong CP problem in QCD. Axions could be produced abundantly in the very early universe, loose a lot of energy during the epoch of inflation and build a cold Bose-Einstein condensate. This evolution mechanism would turn the axion to a very cold dark matter candidate in spite of its small mass.

2.2.2. Warm dark matter (WDM)

The sterile neutrino has been proposed to explain the observed excess of oscillations between muon and and electron neutrinos and antineutrinos [18, and references therein]. Depending on the production mechanism in the early universe, the sterile neutrino could be a very attractive dark matter candidate, and because of its mass at about the keV range, this dark matter would be warm. In the standard production mechanism [19] active neutrinos are converted into a sterile species via oscillations, which are not in thermal equilibrium. However, with high redshift power spectrum measurement from the Lyman-alpha forest as well as constraints on the maximum decaying rate of the sterile neutrinos from X-ray measurements, this production mechanism can be ruled out [20, 21]. More sophisticated production mechanisms including resonant oscillation alleviate this constraints, and the sterile neutrino is maintained as a valid dark matter candidate [22].

Another often mentioned candidate for WDM is the gravitino, which is the supersymmetric partner of the graviton and can have a mass in the keV range. Depending on the scale of symmetry breaking, the gravitino could be the lightest supersymmetric particle and therefore act as the dark matter particle [23].

2.3. Linear structure formation

In the early universe matter is distributed nearly uniformly with tiny density perturbations initially induced by quantum fluctuations and blown up to macroscopic scale through inflation. These perturbations (or overdensities) are usually parametrised as

$$\delta(\mathbf{x}, t) = \frac{\rho(\mathbf{x}, t) - \rho_b(t)}{\rho_b(t)}. \tag{2.6}$$

where $\rho(\mathbf{x}, t)$ is the density at the comoving coordinate $\mathbf{x}$ and $\rho_\mathrm{b}(t)$ is the background density, described by Eq. (2.5). As long as $\delta \ll 1$, it is possible to describe the evolution of the overdensities analytically via a perturbational approach. In this

section we are looking at the linear order evolution of δ in the Newtonian limit, where the wave length is much smaller than the horizon $d_H(t) = H^{-1}(t)$.

2.3.1. Particle motion

As a first step we are deriving the motion of a non relativistic particle of mass m under a smooth gravitational potential in an expanding universe (see [25] for more details). The proper position (in physical space) as well as the proper velocity are given by

$$\mathbf{r} = a\mathbf{x}, \qquad \mathbf{u} = a\dot{\mathbf{x}} + \dot{a}\mathbf{x} = \mathbf{v}_\mathrm{p} + \mathbf{v}_\mathrm{h}, \tag{2.7}$$

where a is the scale factor defined in (2.2). The proper velocity $\mathbf{u}$ is the sum of the peculiar velocity $\mathbf{v}_\mathrm{p}$ and the Hubble flow $\mathbf{v}_\mathrm{h}$. Under the influence of a gravitational potential $\Phi(\mathbf{x}, t)$ the motion of the particle is described by the Lagrangian $L = m\mathbf{u}^2/2 - m\Phi$. A canonical transformation $L \to L - d\psi/dt$ with $\psi = ma\dot{a}\mathbf{x}^2/2$ then leads to

$$L = \frac{1}{2}ma^2\dot{\mathbf{x}}^2 - m\phi(\mathbf{x}, t), \qquad \phi = \Phi + \frac{1}{2}a\ddot{a}\mathbf{x}^2. \tag{2.8}$$

Applying the usual recipe of the Legendre transformation, the corresponding Hamiltonian H is found to be

$$H = \frac{1}{a^2}\frac{\mathbf{p}^2}{2m} + m\phi(\mathbf{x}, t), \tag{2.9}$$

where $\mathbf{p} = ma^2\dot{\mathbf{x}}$ is the comoving momentum.

The new definition for the potential ϕ still has to be set into context. Using the Poisson equation $\nabla^2\Phi = 4\pi G\rho$ as well as the Friedmann equations (2.3, 2.4) it is straight forward to derive the relation

$$\frac{1}{a^2}\nabla^2_\mathbf{x}\phi(\mathbf{x}, t) = 4\pi G\rho_b(t)\delta(\mathbf{x}, t), \tag{2.10}$$

which we will call the modified Poisson equation.

2.3.2. Collisionless Boltzmann equation

In the last section we described the motion of a single particle in expanding space and under the influence of an external potential $\phi(\mathbf{x}, t)$, which on its part is coupled to the overdensity field $\delta(\mathbf{x}, t)$ via the modified Poisson equation (2.10). In order to get the evolution of the whole fluid, we will use the notion of the phase-space density $f(\mathbf{x}, \mathbf{p}, t)$, defined as $dN = f(\mathbf{x}, \mathbf{p}, t)d^3xd^3p$, where N is the number of particles inside the finite phase space volume V. An integration of the phase space density over velocity space leads to the real space proper density

$$\rho(\mathbf{x}, t) = mn(\mathbf{x}, t) = \frac{m}{a^3}\int d^3p f(\mathbf{x}, \mathbf{p}, t) = \frac{\rho_{b0}}{a^3}\left[1 + \delta(\mathbf{x}, t)\right]. \tag{2.11}$$

Here we have used the assumption of a non relativistic fluid obeying the relation $\rho_b(t) = \rho_{b0}a^{-3}$, where ρ_{b0} is the background density at redshift zero. At the same time we have defined the proper number density n of a non relativistic fluid.

Liouville's theorem states that the phase space density of a collisionless fluid is conserved over time. Since dark matter is a nearly perfect collisionless system (with negligible cross section and smoothly varying potential), its phase space density is a constant of time and in the Newtonian limit we can write

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{f, H\} = 0, \qquad \{f, H\} = \frac{\partial f}{\partial \mathbf{x}} \cdot \frac{\partial H}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}} \cdot \frac{\partial H}{\partial \mathbf{x}}. \tag{2.12}$$

Using the Hamiltonain derived in (2.9) we get what is called the collisionless Boltzmann (or Vlasov) equation:

$$\frac{\partial f}{\partial t} + \frac{1}{a^2}\frac{\mathbf{p}}{m} \cdot \nabla_{\mathbf{x}} f - m\nabla_{\mathbf{x}}\phi \cdot \frac{\partial f}{\partial \mathbf{p}} = 0, \tag{2.13}$$

which describes the evolution of a collisionless fluid under the effect of self gravity. Due to the nonlinearity of the Boltzmann equation, general solutions cannot be found and we therefore have to do some simplification. In a first step we will derive the continuity and the momentum equation which can then be combined to find an equation for the overdensity evolution. To do so, we define the local streaming velocity and the velocity stress tensor:

$$\langle \mathbf{v} \rangle = \frac{1}{na^3} \int d^3p \frac{\mathbf{p}}{ma} f, \qquad \langle v_i v_j \rangle = \frac{1}{na^3} \int d^3p \frac{p_i p_j}{(ma)^2} f. \tag{2.14}$$

The continuity equation can now be obtained by simply integrating Eq. (2.13) over momentum space. Using Eq. (2.14) and ignoring the surface term leads to

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla_{\mathbf{x}} \cdot [\langle \mathbf{v} \rangle (1 + \delta)] = 0. \tag{2.15}$$

The first momentum equations are obtained by multiplying Eq. (2.13) with p_i and integrating over momentum space:

$$\frac{\partial}{\partial t} [a\langle v_i \rangle (1 + \delta)] + \partial_i \phi (1 + \delta) + \partial^j [\langle v_i v_j \rangle (1 + \delta)] = 0. \tag{2.16}$$

Taking the time derivative of this relation and substituting it into Eq. (2.15) leads to

$$\frac{\partial^2 \delta}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial \delta}{\partial t} = \frac{1}{a^2}\partial_i \left[(1 + \delta)\partial^i \phi\right] + \frac{1}{a^2}\partial^i \partial^j \left[\langle v_i v_j \rangle (1 + \delta)\right], \tag{2.17}$$

which gives a description of the evolution of density perturbations.

Approximate solutions for $\delta \ll 1$ can be found by linearizing the system of equations (2.17, 2.10). In this regime, the velocity shear is negligible and the stress tensor becomes diagonal. The diagonal terms are simply given by $\sigma/\sqrt{3}$, where σ is the three dimensional initial fluid dispersion. The linearisation then leads to

$$\frac{\partial^2 \delta}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial \delta}{\partial t} = \frac{\sigma^2}{a^2}\nabla^2_{\mathbf{x}} \delta + 4\pi G \rho_{\mathrm{b}} \delta. \tag{2.18}$$

We can now apply a Fourier transform of the form

$$\delta(\mathbf{x}, t) = \int d^3\mathbf{k} \delta_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}} \tag{2.19}$$

to obtain the ordinary linear differential equation

$$\frac{d^2\delta_{\mathbf{k}}}{dt^2} + 2\frac{\dot{a}}{a}\frac{d\delta_{\mathbf{k}}}{dt} = \left[4\pi G\rho_{\mathrm{b}}(t) - \frac{\sigma^2(t)k^2}{a^2}\right]\delta_{\mathbf{k}}. \tag{2.20}$$

Eq. (2.20) describes the behavior of linear perturbations in an idealised universe containing one single fluid. Considering a more realistic approach of many different cosmic fluids (such as non realistic dark matter, baryons, radiation etc), it is straight forward to extend the derivation above to get a set of equations of the form

$$\frac{d^2\delta_A}{dt^2} + 2\frac{\dot{a}}{a}\frac{d\delta_A}{dt} = \left[4\pi G\rho_A(t) - \frac{\sigma_A^2(t)k^2}{a^2}\right]\delta_A + 4\pi G\sum_B \rho_B(t)\delta_B, \tag{2.21}$$

where the subscript A denotes the fluid of interest, whose perturbations are coupled to the perturbations of all the other components of the universe, subscribed with B.

The second term in the square brackets of Eqs. (2.20) and (2.21) becomes only important at small scales (large k) and will be neglected for the moment. A detailed study of the case when this term becomes dominant is given in section 2.3.3.

We are now examining solutions of the equations of perturbations for the different cosmological epochs in order to get a qualitative picture of dark matter structure formation. At early times and scales well below the Hubble radius, the dominant radiation component does not cluster and the coupling in Eq. (2.21) disappears [see 26, for more details]. The perturbations of dark matter then behave as a single non relativistic fluid in a radiation dominated universe (essentially Eq. (2.20) with $\rho_{\mathrm{b}} = \rho_{\mathrm{m}}$). Following Mo et al. [24] we define $\zeta = \rho_m/\rho_r = a/a_{\mathrm{eq}}$, where a_{eq} is the scale factor at matter-radiation equality. Using the Friedmann equations and ignoring the subdominant cosmological constant, we can then rewrite Eq. (2.20) to get

$$2\zeta(1+\zeta)\frac{d^2\delta_{\mathbf{k}}}{d\zeta^2} + (2+3\zeta)\frac{d\delta_{\mathbf{k}}}{d\zeta} = 3\delta_{\mathbf{k}}. \tag{2.22}$$

The growing part of the solution is easily found to be

$$\delta_+ \equiv D(a) \propto 1 + \frac{3}{2}\left(\frac{a}{a_{\mathrm{eq}}}\right). \tag{2.23}$$

In the epoch of radiation domination $\zeta < 1$, and the perturbations stay roughly constant. This is called Mészáros effect. In the epoch of matter domination $\zeta > 1$, and the perturbations are linearly growing with respect to the scale factor.

In the very late universe the cosmological constant becomes important for structure formation. The growing solution of the equation of perturbations must then be calculated numerically. Here we just give the approximate solution from Caroll et al. [27]:

$$\delta_+ \equiv D(a) = g(a)a, \tag{2.24}$$

$$g(a) = \frac{5}{2}\Omega_{\mathrm{m}}\left[\Omega_{\mathrm{m}}^{\frac{4}{7}} - \Omega_\Lambda + \left(1+\frac{\Omega_{\mathrm{m}}}{2}\right)\left(1+\frac{\Omega_\Lambda}{70}\right)\right]^{-1}, \tag{2.25}$$

where

$$\Omega_{\rm m} = \frac{\Omega_{\rm m,0}}{a^3 H(a)}, \qquad \Omega_\Lambda = \frac{\Omega_{\Lambda,0}}{H(a)}, \tag{2.26}$$

and $H(a) = \dot{a}/a$ is the Hubble parameter obtained by the first Friedmann equation (2.3). The subscript 0 denotes values at redshift zero.

During the cosmic history, significant formation of structures only starts after matter-radiation equality, and the growth holds on until today, although it slows down with the increasing dominance of the cosmological constant.

2.3.3. The Jeans criterion and the free streaming scale

Up to now, we have ignored the effect of the velocity dispersion on structure formation. This is a good approximation as long as the scales considered are large enough. On smaller scales however, the dispersion has the effect of damping the growth of structures or suppressing it completely. Equation (2.20) only has growing mode solutions if the term on the right hand side of the equality stays above zero:

$$\left[4\pi G\rho_{\rm b}(t) - \frac{\sigma^2(t)k^2}{a^2}\right] \geq 0. \tag{2.27}$$

If it becomes negative, the density perturbations start to oscillate and no structures can grow. We can therefore define the physical Jeans length

$$\lambda_J(t) = a\frac{2\pi}{k_J} \equiv \sqrt{\frac{\pi\sigma^2(t)}{G\rho_{\rm b}(t)}}, \tag{2.28}$$

which denotes the limit between stability and collapse. A more intuitive scale in terms of haloes is the Jeans mass, which is defined as

$$M_J \equiv \frac{4\pi}{3}\rho_{\rm m}\left(\frac{\lambda_J}{2}\right)^3. \tag{2.29}$$

Note that while the Jeans length depends on the background density, which includes all cosmic components and goes as $\rho_{\rm b} \propto a^{-4}$ for $a < a_{\rm eq}$, the conversion between Jeans length and Jeans mass depends on the dark matter density with the behvior $\rho_{\rm m} \propto a^{-4}$ for $a < a_{\rm nr}$ and $\rho_{\rm m} \propto a^{-3}$ for $a > a_{\rm nr}$.

We can now qualitatively trace the evolution of the Jeans criterion through cosmic history. In the very early universe during in the epoch of radiation domination, even the dark matter particles are highly relativistic with a dispersion of the order $\sigma(t) \sim 1$. During this epoch the Jeans mass is growing as $M_J \propto a^2$. Since the dark matter particle is very heavy, it becomes non relativistic while the universe is still dominated by radiation. In this epoch $\sigma \propto a^{-1}$ and M_J is constant. Later on, in the epoch of matter domination, the Jeans mass is shrinking again with $M_J \propto a^{-3/2}$. (For small perturbations that already entered the horizon at $a < a_{\rm nr}$ the dark matter perturbations decouple from radiation and the Jeans mass decreases as $M_J \propto a^{-3/2}$, already at $a_{\rm nr} \leq a \leq a_{\rm eq}$.) The behavior of the Jeans mass is summarised in Fig. 2.1, where the red colored region represents scales with no growth of perturbations.

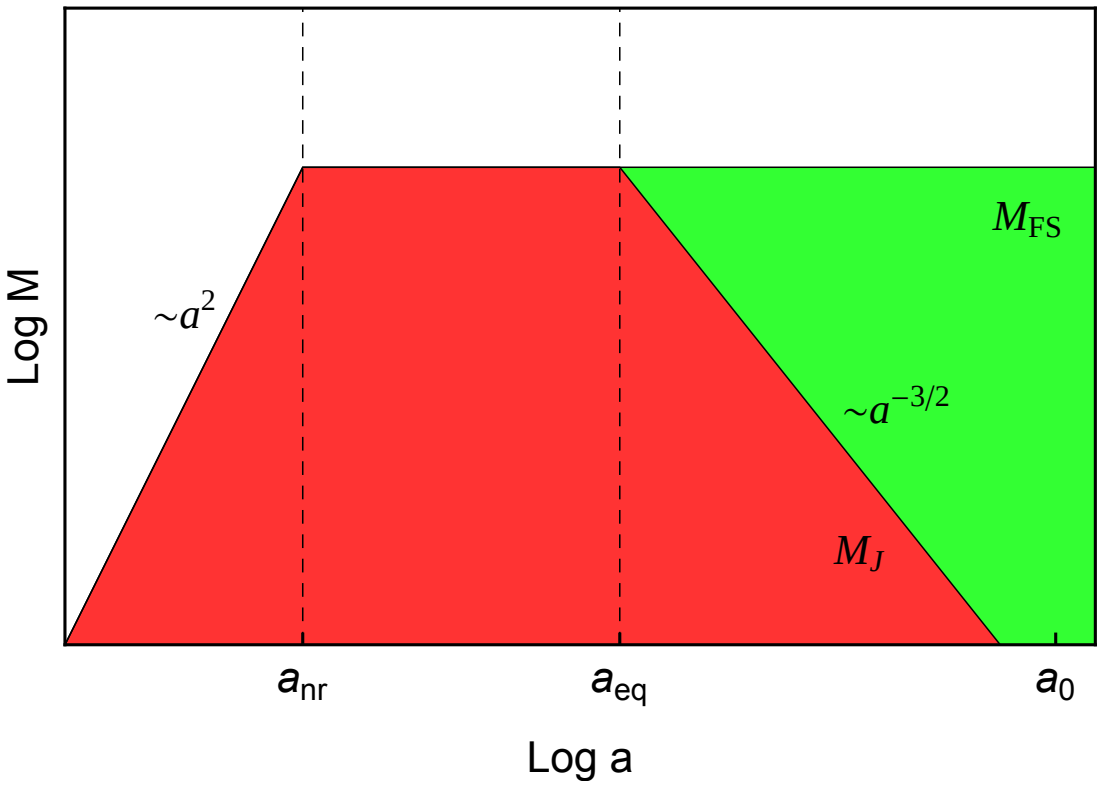

Figure 2.1. Schematic representation of the Jeans mass M_J and the free streaming mass $M_{\rm FS}$ as they evolve through cosmic history. In the red colored region the density perturbations are not growing because they are Jeans stable. The green colored region is Jeans unstable but the perturbations are completely wiped out due to the velocity free streaming.

At scales below the Jeans length, particles are free streaming out of overdensities and completely destroy all density perturbations. Therefore, there is no linear growth of structures below the Jeans scale of matter-radiation equality, although M_J is shrinking dramatically afterwards. The free streaming length is usually defined as

$$\lambda_{\rm FS} = \int_0^{t_{\rm eq}} \frac{\sigma(t)}{a} dt, \tag{2.30}$$

which is the streaming distance of a dark matter particle from the big bang until $t_{\rm eq}$, when the modes become Jeans unstable. Since $\sigma \sim c$ for $t \leq t_{\rm nr}$ and $\sigma \sim \sigma_0/a$ for $t \geq t_{\rm nr}$ this integral can be split into two parts, yielding the approximate solution

$$\lambda_{\rm FS} \sim \frac{2ct_{\rm nr}}{a_{\rm nr}} \left[1 + \log\left(\frac{a_{\rm eq}}{a_{\rm nr}}\right)\right], \tag{2.31}$$

where we have used $a \propto t^{1/2}$ for the epoch of radiation domination.

This definition of the free streaming length happens to be approximately equivalent to calculating the Jeans length at matter-radiation equality. Using Eq. (2.28) as well as the approximative relation $\rho_{\rm b} = \rho_0/a^3$ up to matter-radiation equality gives

$$\lambda_J^{\rm eq} \sim \sqrt{\frac{\pi\sigma_0^2}{G\rho_0 a_{\rm eq}}} \sim \frac{90{\rm Gyr}}{\sqrt{a_{\rm eq}}}\sigma_0 \tag{2.32}$$

in comoving units. An estimation of the free streaming length leads to

$$\lambda_{\rm FS} \sim \frac{2ct_{\rm nr}}{a_{\rm nr}} \sim \frac{2t_0}{\sqrt{a_{eq}}}\sigma_0 \sim \frac{120{\rm Gyr}}{\sqrt{a_{eq}}}\sigma_0. \tag{2.33}$$

Here we have used the approximation $c \sim \sigma_0/a_{\rm nr}$ and we have set to three the logarithm in Eq. (2.31). Both relations have a linear dependence on the particle dispersion and are of roughly the same size. This confirms the qualitative picture given in Fig. 2.1 .

2.4. Statistical description of the density field

The overdensity field $\delta(\mathbf{x}, t)$ is a complicated and dynamical scalar field, which governs structure formation and is traced by galaxies. For a detailed comparison with the observed galaxy distribution, it is therefore crucial to apply a statistical approach to the overdensity field.

2.4.1. Basic concepts

The simplest and most straight forward statistical tool to describe the distribution of matter in the universe is given by the two point correlation function

$$\xi(r) = \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle. \tag{2.34}$$

Note that $\xi(r)$ only depends on the norm of $\mathbf{r}$, which is a direct consequence of the statistical homogeneity of space. The ensemble average $\langle\cdot\rangle$ can be understood as the mean over an infinite amount of independent realisations of the field (or equivalently as the mean over all positions in space). In Fourier space the two point correlation becomes

$$\langle \delta(\mathbf{k})\delta(\mathbf{q})\rangle = \frac{1}{(2\pi)^6}\int d^3\mathbf{x}d^3\mathbf{r}\langle\delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle e^{-i[(\mathbf{k}+\mathbf{q})\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{r}]}. \tag{2.35}$$

Here we have used the Fourier transform of the density perturbation defined in Eq. (2.19), and we changed the notation from $\delta_{\mathbf{k}}$ to $\delta(\mathbf{k})$. With the Fourier transform of the Dirac delta

$$\delta_D(\mathbf{k}+\mathbf{q}) = \frac{1}{(2\pi)^3}\int d^3\mathbf{x}e^{-i(\mathbf{k}+\mathbf{q})\cdot\mathbf{x}}, \tag{2.36}$$

relation (2.35) can be written as

$$\langle \delta(\mathbf{k})\delta(\mathbf{q})\rangle = \delta_D(\mathbf{k}+\mathbf{q})\frac{1}{(2\pi)^3}\int d^3\mathbf{r}\xi(r)e^{-i\mathbf{q}\cdot\mathbf{r}} =: \delta_D(\mathbf{k}+\mathbf{q})P(k), \tag{2.37}$$

where we have defined the power spectrum $P(k)$, which simply is the Fourier transform of the correlation function:

$$\xi(r) = \int d^3\mathbf{k}P(k)e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{2.38}$$

For a comparison with observations or N-body simulations, it is useful to smooth the overdensity field by integrating over a window function

$$\delta(\mathbf{x}, R) = \int d^3\mathbf{x}'\delta(\mathbf{x}')W(|\mathbf{x} - \mathbf{x}'|, R). \tag{2.39}$$

The most natural choice for the window function is a real space tophat function of the form $W(x, R) = 3\Theta(R-x)/4\pi R^3$, where Θ denotes the Heaviside step function. Other popular choices are the Gaussian window $W(x, R) = \exp(x^2/2R^2)/(2\pi R^2)^{3/2}$ or the k-space tophat window $W(y) = (\sin y - y\cos y)/(2\pi^2 R^3 y^3)$, where $y = x/R$.

2.4.2. Gaussian random field

The model of inflation predicts that the universe is not completely homogeneous, but that there are tiny quantum fluctuations of space, which are blown up from initially microscopic scales through exponential expansion of space [28]. These quantum fluctuations are the seeds of the density perturbations and of all galaxies in the universe. A basic prediction of inflation (at least of the standard single field inflation) is that the overdensities are distributed as a Gaussian random field. For the linear overdensity field smoothed at a scale R, the distribution can therefore be written as

$$P(\delta|R)d\delta = \frac{1}{\sqrt{2\pi\sigma^2(R)}} \exp\left(-\frac{\delta^2}{2\sigma^2(R)}\right)d\delta, \tag{2.40}$$

where the variance $\sigma^2(R)$ is given by the ensemble average of the density squared:

$$\sigma^2(R) = \langle\delta(\mathbf{x}, R)^2\rangle = \frac{1}{(2\pi^3)} \int d^3\mathbf{k}P(k)W^2(k; R). \tag{2.41}$$

Up to now, neither CMB measurements nor large scale galaxy surveys have found any deviation from this Gaussian distribution in the linear field [1]. The search for non-Gaussianities is nevertheless a very active field of research, since the discovery of non-Gaussian features would permit to constrain inflation and give an insight into the nature of the very early universe. The Gaussian nature of the field gets lost as soon as the growth of the perturbations becomes nonlinear. Already second order contributions to the perturbed field induce non-Gaussian features, and the search of primordial non-Gaussianities is therefore restricted to the linear scales.

Another basic prediction from single field inflation is a nearly scale-invariant primordial power spectrum of the form

$$P(k) \propto k^{n_s}T^2(k), \tag{2.42}$$

where the spectral index n_s is approximately one. The transfer function $T(k)$ describes the suppression of the perturbations after inflation and catches the complicated interplay of sub- and super-horizon evolution of density perturbations during the radiation dominated epoch. It has to be determined by numerically solving the relativistic Boltzmann equation. In ΛCDM, a good fit to such a numerical solution is given by

$$T(k) = \frac{\log(1 + 2.34q)}{2.34q}\left[1 + 3.89q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4\right]^{-1/4}, \tag{2.43}$$

where $q = k/\Omega_m h$ [29]. The suppression of the transfer function is largest at small scales, indicating the absence of growth below the horizon during the epoch radiation domination.

For a ΛWDM cosmology, equation (2.42) has to be multiplied by the fitting function [30]

$$T_{\mathrm{WDM}}(k) = \left[1 + (\alpha k)^{2\mu}\right]^{-5/\nu}, \tag{2.44}$$

where $\nu = 1.12$ and

$$\alpha = 0.049 \left[\frac{m_{\mathrm{WDM}}}{\mathrm{keV}}\right]^{-1.11} \left[\frac{\Omega_{\mathrm{WDM}}}{0.25}\right]^{0.11} \left[\frac{h}{0.7}\right]^{1.22} \mathrm{Mpc/h}. \tag{2.45}$$

This additional truncation in the ΛWDM power spectrum induced by T_{WDM} comes from the fact that the particle free streaming suppresses power at small scales.

The power spectrum together with the Gaussian nature of the perturbations gives a statistically complete description of the overdensity field, which can be used as a starting point to understand the complicated and non linear behavior of structure formation.

2.5. Nonlinear structure formation

The linear perturbational approach described in the last section is approximately right for small overdensities, but it breaks down completely as soon as $\delta \geq 1$. In this nonlinear regime the system of equations (2.10, 2.17) has to be solved directly, something that can only be done with numerical simulations. We will now summarise the most popular techniques to perform numerical simulation, followed by a discussion on basic analytical and semi-analytical approaches.

2.5.1. Numerical simulations

The common techniques of numerical structure formation are generally based on a large number of N hypothetical particles representing fluid elements, each moving with respect to its conjoint force field as

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{u}_i, \qquad \mathbf{a}_i = \frac{d\mathbf{u}_i}{dt} = -\nabla\Phi(\mathbf{r}_i). \tag{2.46}$$

This approach is equivalent to solving the system of equations (2.10, 2.17), at least in the limit of $N \to \infty$. In reality N is limited by computational power and memory, and the simulation is only accurate down to a certain resolution limit.

The numerical integration is usually done by the 'leap frog' scheme, where the particles first get drifted for half a time step in real space

$$\mathbf{r}_i(t_{n+1/2}) = \mathbf{r}_i(t_n) + \mathbf{u}_i(t_n))\frac{\Delta t}{2}, \tag{2.47}$$

followed by a recalculation of the forces for the new particle positions. In a next step the particle velocities are kicked for a full time step

$$\mathbf{u}_i(t_{n+1}) = \mathbf{u}_i(t_n) + \mathbf{a}_i(t_{n+1/2})\frac{\Delta t}{2}, \tag{2.48}$$

before another drift for half a time step

$$\mathbf{r}_i(t_{n+1}) = \mathbf{r}_i(t_n) + [\mathbf{u}_i(t_n) + \mathbf{u}_i(t_{n+1}))] \frac{\Delta t}{2} \tag{2.49}$$

completes one cycle. This procedure is repeated over and over until the end of the numerical integration. The 'leapfrog' algorithm is very popular because of its simplicity and because it is an energy conserving integrator.

The computationally expensive part of the leapfrog algorithm is the force calculation, and there are many different techniques for doing this efficiently. The most common ones are summarised in the following:

Particle-particle (PP) algorithm: The most straight forward way of calculating the forces is direct summation over all particle contributions. This procedure is however only a valid option for systems with small N. For larger setups it becomes tremendously expensive, since the computational costs grow with $\mathcal{O}(N^2)$.

Treecode: A more sophisticated technique, initially introduced by Barnes & Hut [31], consists of grouping more distant particles together and calculating their force contribution with one single multipole expansion. The division into groups is usually done by an octree, which divides the whole simulation box into eight equally sized sub-cells, that are again divided, as long as there is more than one particle in a cell. A fixed opening angle then decides how many particles can be treated within a single multipole expansion. In the treecode algorithm the costs of the force calculation drops down to $\mathcal{O}(N \log N)$, which makes it very popular for large N-body simulations.

Particle-mesh (PM) algorithm: In this technique a fine-meshed grid is spanned over the whole simulation box and particles are transferred to the grid points, turning the particle description to a discrete density scalar field. The grid is then transformed to Fourier space with a fast Fourier transform (FFT) method, where the Poisson equation simply becomes $-k^2\Phi_{\mathbf{k}} = 4\pi\rho_{\mathbf{k}}$ and the force is given by $\mathbf{a}_{\mathbf{k}} = -i\phi_{\mathbf{k}}\mathbf{k}$. An inverse Fourier transformation of $\mathbf{a}_{\mathbf{k}}$ and an interpolation back to the particle positions then gives the forces in real space for every particle. This method is very fast with computational costs of $\mathcal{O}(N)$, but it has the drawback of being limited in resolution by the size of one grid cell.

Particle-particle-particle-mesh (PPPM) algorithm: This hybrid method combines the advantages of both tree and PM-code and is often used for cosmological simulations. The long range forces are thereby calculated on a grid and the treecode is only used below a certain length threshold.

Additionally to sophisticated force calculations, numerical integrators generally include adaptive time stepping, where the force calculation is done more often for particles in a strong gravitational field. Two particles coming together very close need very small time steps to be treated accurately. An artificial force softening prevents a divergence of the force and the freezing of the code. The softening induces a resolution length, below which simulations can not be trusted anymore.

In this thesis we used the gravity code PKDGRAV, a fully parallelised treecode with multipole expansion up to seventh order and adaptive time stepping. PKDGRAV was written by Joachim Stadel [32].

2.5.2. Approximative analytical approaches

The nonlinear structure formation can be approximately described with analytical and semi-analytical methods. This is generally done by extrapolating from the linear density field using some criteria for an instantaneous halo collapse.

The most basic approach, the Press-Schechter (PS) model [33], is based on spherical collapse, where a tophat spherical overdensity, initially in the Hubble flow, gets amplified, decouples, and turns around to collapse and virialise. The time of collapse corresponds to a linear overdensity growth of $\delta_c = 1.686$ and this value is therefore taken as a collapse threshold. In the Press-Schechter model, the density field is evolved linearly well above the usual limitation $\delta \ll 1$. As soon as the field exceeds the threshold ($\delta > \delta_c$) a halo of the corresponding smoothing scale R is thought to form and virialise immediately. The cumulative probability of having a region of size R above the threshold is given by

$$F(R) = \int_{\delta_c}^{\infty} P(\delta|R) d\delta = \frac{1}{2}\mathrm{erfc}\left[\sqrt{\frac{\nu}{2}}\right], \tag{2.50}$$

where $\nu = [\delta_c/\sigma(R)]^2$ and $P(\delta|R)$ is the Gaussian distribution (2.40). Since in ΛCDM $\sigma(R)$ is a monotonically decreasing function of R, the probability of a patch to be above the threshold δ_c is larger for smaller R. A qualitative picture of the Press-Schechter model is given in Fig. 2.2. The colored curves represent the Gaussian distribution of overdensities with different values for σ, or equivalently different smoothing scale R. The flatter the curve (large σ) the smaller the smoothing scale we look at. The filled region at $\delta > \delta_c$ corresponds to the cumulative probability $F(R)$ and grows with decreasing scale R. There is hence a larger fraction of collapsed volume elements at small scales, which is a consequence of hierarchical structure formation. The time evolution of the PS model is encapsulated in $\sigma(R) = D(a)\sigma_0(R)$, what means that the Gaussian curve is flattening with time and $F(R)$ is growing. An analog description is setting σ constant and letting the threshold evolve with time ($\delta_c \propto 1/D(a)$; the dashed line in Fig. 2.2 is then moving to the left of the plot until the displayed value at redshift zero). In the hierarchical picture of ΛCDM, the variance diverges for $R \to 0$ and all mass should be collapsed. This is however not the case in Eq. (2.50) where $F(0) = 1/2$. Press & Schechter [33] corrected this inconsistency by simply multiplying (2.50) with a factor of two, however without giving a convincing physical justification. With this additional 'fudge' factor, the number density of haloes with mass $M = 4\pi\rho_\mathrm{b}R^3/3$ is given by

$$\frac{dn}{d\log M} = \frac{\rho_\mathrm{b}}{M}\left|\frac{dF(M)}{d\log M}\right| = -\frac{1}{2}\frac{\rho_\mathrm{b}}{M} f_\mathrm{PS}(\nu)\frac{d\log\sigma^2}{d\log M},\ f_\mathrm{PS}(\nu) = \sqrt{\frac{2\nu}{\pi}}\exp\left[-\frac{\nu}{2}\right]. \tag{2.51}$$

In spite of the very simplistic approach, the PS halo mass function is in reasonable agreement with ΛCDM (see Fig. 2.3).

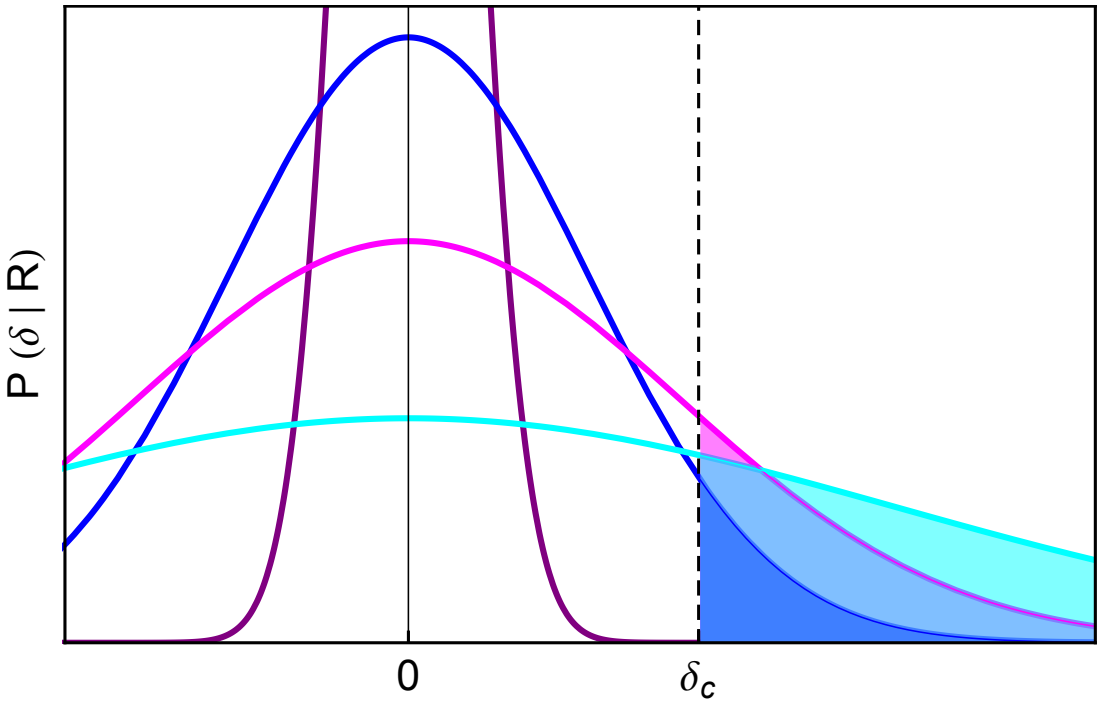

Figure 2.2. Gaussian distribution function of the density perturbations for different variance $\sigma^2(R)$. The flatter the curve, the higher the variance and the smaller the smoothing scale R. The collored area above δ_c corresponds to the cumulative probability $F(R)$.

The main conceptual weakness of the PS model is the fact that it does not account for the so called cloud-in-cloud problem, meaning that an underdense patch within a larger overdense region is not counted as collapsed, even if the larger region is above the threshold. The excursion set approach or extended Press-Schechter (EPS) model solves the cloud-in-cloud problem by following every overdensity patch from the large, homogeneous scales down to smaller and smaller smoothing scales (or equivalentely from very high redshift to redshift zero). As soon as a patch first crosses the threshold the patch is counted as virialised and a second crossing is not allowed. The EPS approach leads to a halo mass function of Eq. (2.51), but this time without having to introduce the 'fudge' factor of the initial PS approach. A detailed derivation of the excursion set approach can be found in Zentner [34].

The real advantage of the EPS approach is that it gives a method for further investigations and more detailed modeling. It becomes for example possible to replace the spherical collapse by the more realistic ellipsoidal collapse model, where the threshold barrier itself is scale dependent [35]. A often used approach inspired by ellipsoidal collapse and adapted to numerical simulations is the Sheth-Tormen (ST) mass function

$$\frac{dn}{d\log M} = -\frac{1}{2}\frac{\rho_\mathrm{b}}{M} f_\mathrm{ST}(\nu)\frac{d\log\sigma^2}{d\log M}, \quad f_\mathrm{ST}(\nu) = \sqrt{\frac{2q\nu}{\pi}}\left[1+(q\nu)^{-p}\right]\exp\left[-\frac{q\nu}{2}\right], \quad (2.52)$$

which is a very accurate and physically well motivated fit to CDM simulations (see Fig. 2.3 for a comparison with simulations).

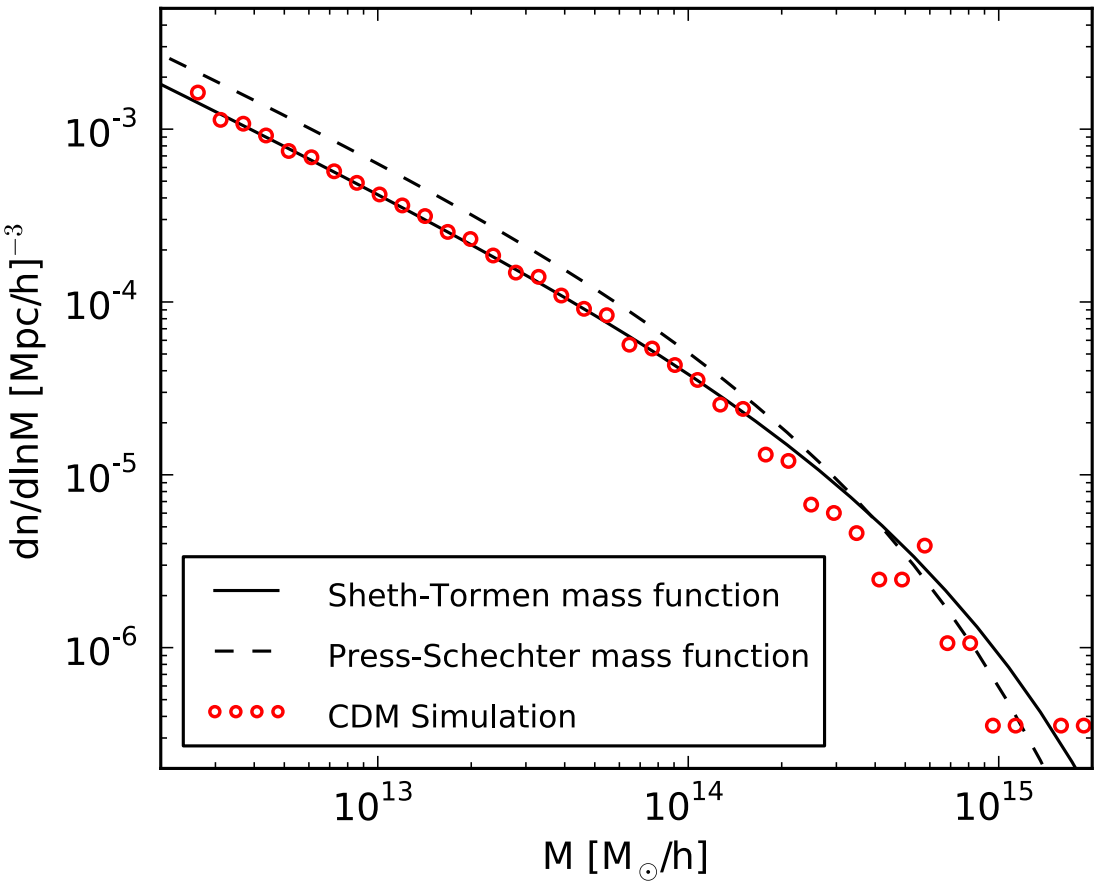

Figure 2.3. Halo mass function from the ST approach (solid line) and the PS approach (dashed line) in comparison with simulations (red dots).

Next to the halo mass function the EPS formalism can also be used to compute characteristic quantities of structure formation, like the halo bias, halo merger trees, or void statistics. In spite of its simplicity it gives a good qualitative picture of the nature of nonlinear structure formation. Accurate quantitative results can however only be obtained with fully consistent numerical simulations.

3

Paper I: CDM microhaloes

In this first central part of the thesis we are considering the classical ΛCDM picture with a 100 GeV neutralino as dark matter particle candidate. Due to its large mass and small cross section, the neutralino becomes nonrelativistic and decouples from the plasma at very early times. Consequently, the free streaming length is extremely small, leading to hierarchical structure formation over an enormous range of scales.

The smallest dark matter halo is expected to be around the free streaming mass, which can be approximately determined with the relation (2.33). Taking the velocity dispersion at decoupling $\sigma_d \sim (3k_B T_d/m)^{1/2}$ as well as the assumptions $m \sim 100$ GeV and $T_d \sim a_d^{-2}$ for particle mass and decoupling temperature, one gets a free streaming mass scale of roughly $M_{\rm FS} \sim 10^{-6} M_\odot$. The precise effect of free streaming on the initial matter power spectrum depends on the characteristics of the WIMP as well as the cosmic history, and it can only be calculated by solving the relativistic Boltzmann equation [36, 37].

The number density of the whole hierarchy of structures is given by the mass function, and although numerical simulations do not come close to a comprable resolution, there are reasons to believe that the mass function behaves as a power law of the form

$$\frac{dn}{d\log M} \sim M^{-1} \tag{3.1}$$

all the way up to $M \sim 10^{-6} M_\odot$. First of all the EPS formalism predicts a power law behavior at small masses and second numerical simulation of ΛCDM show no significant deviation from a power law within the range they can challenge.

If this extrapolation of the mass function is correct, there must be an enormous amount of microhaloes in the universe. Considering our milky way halo, about 10^{13} microhaloes in the mass range $[10^{-6}, 10^{-5}] M_\odot$ should be found within the virial radius [38]. The hierarchical nature of CDM structure formation predicts the microhaloes to be the structures that form first and to have the highest concentrations. They are therefore the most probable to survive the tidal forces within the milky

way halo and are potential interesting objects for direct and indirect dark matter detection experiments.

There is a major ongoing experimental effort to proof the existence of neutralino dark matter either via direct detection by measuring recoil energies or via indirect detection by observing a gamma ray signal coming from dark matter self annihilation (see section 2.2.1 for more details). For the interpretation of these experiments, it is crucial to have an idea of the expected dark matter distribution in the milky way halo and in the solar neighborhood. Predicting the local dark matter distribution around the solar neighborhood is however a very difficult task, since this is a highly nonlinear regime and cosmological N-body simulations have by far not enough resolution to challenge these small scales.

In the following manuscript we are trying to get an idea of the local dark matter distribution by using basic analytical approximation as well as toy model simulations of substructures within the milky way galaxy. We find that although the milky way halo is populated by a large number of substructures spreading over an enormous range of scales, the local dark matter distribution is expected to be very smooth with a close to Maxwellian velocity distribution.

The following paper has been published 2010 in Physical Review D [39].

Impact of Dark Matter Microhalos on Signatures for Direct and Indirect Detection

Aurel Schneider[1], Lawrence Krauss[1,2] and Ben Moore[1]

[1] *Institute for Theoretical Physics, University of Zurich, Zurich, Switzerland* [2] *School of Earth and Space Exploration and Department of Physics, Arizona State University, PO Box 871404, Tempe, AZ 85287;*

(Dated: August 19, 2010)

Detecting dark matter as it streams through detectors on Earth relies on knowledge of its phase space density on a scale comparable to the size of our solar system. Numerical simulations predict that our Galactic halo contains an enormous hierarchy of substructures, streams and caustics, the remnants of the merging hierarchy that began with tiny Earth mass microhalos. If these bound or coherent structures persist until the present time, they could dramatically alter signatures for the detection of weakly interacting elementary particle dark matter (WIMP). Using numerical simulations that follow the coarse grained tidal disruption within the Galactic potential and fine grained heating from stellar encounters, we find that microhalos, streams and caustics have a negligible likelihood of impacting direct detection signatures implying that dark matter constraints derived using simple smooth halo models are relatively robust. We also find that many dense central cusps survive, yielding a small enhancement in the signal for indirect detection experiments.

PACS numbers: 98.80

Introduction

In a ΛCDM dominated universe, structure forms by the hierarchical clustering and merging of small density perturbations [1]. Numerical simulations that follow these processes predict that our Galactic halo should contain a vast hierarchy of surviving substructures - the remnants of the entire halo merger tree [2]. The number density of substructures of a given mass M goes as $n \propto M_{subs}^{-1}$ and they span over 15 decades in mass [3]. The smallest, oldest and most abundant are Earth-mass microhalos with a half mass radius of 10^{-2} pc that formed at $z \simeq 80 - 30$ [4–8]. This minimum mass is modulated by the free streaming velocity which is related to the mass of the neutralino [9, 10].

Simulations of relatively large subhalos suggest that their gravitational interactions with a disk potential, can lead to a destruction of subhalos at distances closer than 30 kpc [11]. Smaller subhalos form earlier, however, with denser cores, and are therefore the most probable dark matter structures to survive gravitational interactions. A second source of fine grained structure to survive are the numerous caustic sheets and folds that form due to the very high initial phase space density of the cold dark matter particles [12]. These are wrapped in a complex way within all the subsequent structures that form, however in the absence of a heating term, the fine grained phase density would be preserved.

With low internal velocity dispersion and high mean density, both the event rate and the characteristic spectrum of energy deposited by dark matter in direct detection experiments [13] could be affected by any features surviving in the phase space distribution of CDM particles. Direct detection experiments are sensitive to the density and velocity distributions of WIMPs on a scale of $\approx 10^{13} m$, the distance the Earth travels over a year. In order to make predictions and exclusion limits, these experiments assume that the dark matter is completely smooth on these scales, with a well mixed Maxwell-Boltzmann velocity distribution [14, 15]. Furthermore, if any such small high-density clumps actually dominate the dark matter distribution in the solar neighborhood, the indirect detection signal due to dark matter annihilation in the galaxy might also be affected.

Since existing N-body simulations of galaxy formation do not have a resolution that goes down to objects with mass as small as $10^{-6} M_{\odot}$, in order to address the question of the survival and impact of such microhalos we need to combine analytical estimates with the results of smaller scale simulations that can resolve such objects.

Previous studies have made analytic [16] and numerical estimates [17] of the disruption timescale of microhaloes as they orbit through the stellar field. Zhao et. al. [16] argued that most of the microhaloes should be completely destroyed by encounters with stars, whilst Goerdt et. al. [17] show that indeed, whilst most of the mass is unbound the dense central cusp may survive intact.

We extend this work by numerical calculating the tidal disruption of microhaloes as they actually orbit through a field of stars, as well as self-consistently including the Galactic halo potential. We calculate the survival statistics of microhaloes using realistic orbital distributions within the disk, allowing us to follow the dynamical structure of the dark matter streams and thus to estimate the fine grained phase space distribution function of WIMPs on scales relevant to dark matter detection experiments.

Microhalo parameters and disruption processes

The dominant processes that can affect microhalos involve gravitational interactions with baryons in the stellar field during the crossing of the disk and also tidal effects of the disk potential during the orbit of the micro-

halo. Unfortunately we cannot account for both effects simultaneously in our simulations as it would require setting up a self-consistent disk with billions of stars. Instead we look separately at the effects of stellar disruption and tidal streaming and we then estimate the combined behavior.

As a substructure halo crosses through a stellar field, high-speed interactions with single stars will heat up the halo distribution, causing it to increase its velocity dispersion and hence its scale size will grow. This process is analogous to galaxy harassment that occurs in clusters [18] and basically has a timescale proportional to the relaxation time of the stellar disk and the time each microhalo spends within the fluctuating potential field of the stars. For an analytical estimation of this process we follow the Goerdt et. al. paper [17].

In the 'distant-tide' approximation [19] the internal energy increase of the microhalo due to a single encounter with a fixed star is given by

$$\delta E(b) = \frac{1}{2}\left(\frac{2GM_*}{b^2 V_{mh}}\right)^2 \frac{2}{3}\langle r^2\rangle, \qquad (1)$$

where b is the impact parameter and V_{mh} is the velocity of the microhalo. Since halos with an early formation time have a low concentration we can set $\langle r\rangle \approx 0.5 r_{vir}$.

One encounter can totally disrupt the microhalo if δE exceeds the binding energy E_b. Since $E_b \approx 0.4 v_{vir}^2$ with $v_{vir}^2 = G m_{vir}/r_{vir}$ [20], the minimal encounter parameter that does not entirely disrupt the microhalo is found to be

$$b_{min} \approx 0.8\left(\frac{GM_* r_{vir}}{V_{mh} v_{vir}}\right)^{1/2}. \qquad (2)$$

We can now define the disruption probability of a microhalo in a stellar field

$$p = \frac{1}{E_b}\int \delta E dN = \int_0^{b_{min}} dN + \frac{1}{E_b}\int_{b_{min}}^{\infty} \delta E dN, \qquad (3)$$

where $dN = 2\pi n b db V_{mh} dt$. Here we have used $\delta E = E_b$ for $b < b_{min}$. Performing the integration leads to a disruption probability of

$$p \approx 4GM_* n_* t\left(\frac{r_{vir}}{v_{vir}}\right) = \frac{2GM_* n_* t}{5H_0\Omega_{m0}^{1/2}}(1+z)^{-3/2}, \qquad (4)$$

where we have used the definition of the virial radius $M_{vir} = \frac{4\pi}{3} r_{vir}^3 200\rho_c$ with $\rho_c = \frac{3H^2}{8\pi G}$. The microhalo is completely destroyed at $p = 1$. Therefore we get the average disruption time

$$t = 250\left(\frac{0.04 M_\odot pc^{-3}}{M_* n_*}\right)\left(\frac{1+z}{61}\right)^{3/2} Myr. \qquad (5)$$

A microhalo with a formation redshift $z \sim 60$ should therefore survive about 250 Myr in a stellar field with a density similar to the one in the solar neighbourhod. However this is only true on average, since one very close encounter can immediately lead to total disruption.

The above estimate does not take into account the internal structure of the microhalos and should therefore only give a very rough estimation of the disruption time. Also it does not allow one to follow the mass loss during the disruption process. We therefore perform a simulation where a microhalo is crossing a periodic box of stars. The box has a length of 50 pc and is filled with randomly distributed stars with the density $\rho = 0.04 M_\odot pc^{-3}$ and the velocity dispersion $\sigma = 50$ km/s. This constellation corresponds to the stellar field in the disk at the solar radius [19]. For simplicity all the stars have the average mass of 0.7 $M_\odot$. The microhalo which is crossing the box at 200 km/s is set up by the halogen-code of Zemp et. al. [21]. Corresponding to the results in [7] it has a mass of $10^{-6} M_\odot$ and the density profile

$$\rho(r) \propto \frac{1}{\left(\frac{r}{r_s}\right)^\gamma \left(1+\left(\frac{r}{r_s}\right)^\alpha\right)^{\frac{\beta-\gamma}{\alpha}}} \qquad (6)$$

with $\alpha = 1$, $\beta = 3$ and $\gamma = 1.2$, as well as a concentration parameter of $c = r_{200}/r_s = 1.6$. The virial radius r_{200} is defined with respect to the background density at $z = 60$, the average formation redshift of a microhalo [4].

In our simulation we find that 50% of the microhalo mass is unbound after 80 Myr of box-crossing (which corresponds to about 40 perpendicular disk passages). After 160 Myr (80 disk passages) even the central core starts to disappear and more than 90% of the microhalo is completely disrupted (see pictures in Table I). At latest after 200 Myr (100 disk passsages) no bound structure is left (see Fig 1). Our simulation gives therefore a slightly shorter disruption time than the simplified analytical estimation of equation (5).

In order to determine what fraction of microhalos survive until the present day, we have to calculate orbital statistics and the distribution of disk crossing times. This can be established by tracing back the orbits of particles in a galactic potential. We use the standard Milky Way model with disk and halo particles set up by the GalactICS code [22] and we select a sample of halo particles in a small box around the position of the sun. The orbits of these particles are followed backwards in time and we find that the average number of disk crossings for these particles is $\bar{c} = 80$ with a standard deviation of $\sigma_c = 43$. The average crossing radius is (not surprisingly) $\overline{R} = 8$ kpc with $\sigma_R = 4$ kpc. The spread of disk crossing events for different particles follows a Maxwell-Boltzmann distribution.

We use this disk crossing distribution combined with the rate of mass loss determined from our numerical study to calculate the survival statistics of microhalos in the vicinity of the sun. Since the timescale for complete disruption in our simulation is equivalent to the average time a microhalo spends in the stellar disk, we conclude that the average microhalo in the vicinity of the sun is just about to be entirely destroyed at the present time

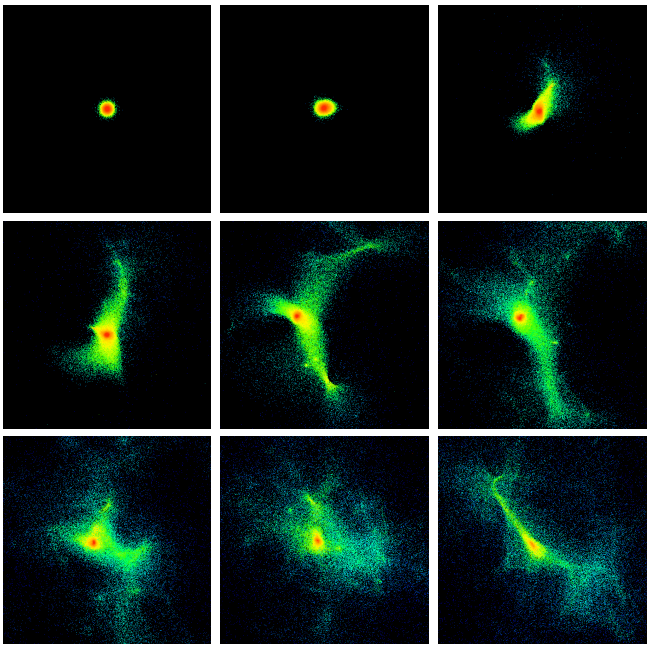

TABLE I: *Microhalo density map at t = 0, 20, 40, 60, 80, 100, 120, 140, 160 Myr (from the upper left to the lower right). The boxlength of the images is 0.38 pc.*

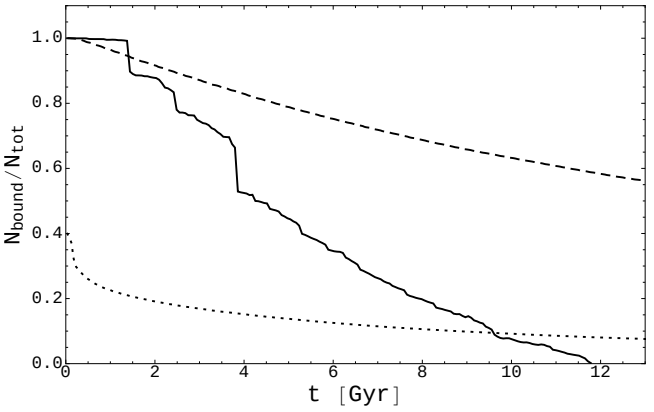

FIG. 1: *Ratio between bound and total mass of a microhalo crossing a stellar field (solid line) and orbiting in a Milky Way potential (dashed, dotted). The solid curve is stretched in order to simulate the effect of disk crossing on an average microhalo with about 80 disk crossings. The dashed and the dotted curve correspond to a microhalo that spent 0 Myr respectively 80 Myr in the stellar field before orbiting around the galactic potential. The dominant mass loss is coming from stellar interaction until complete disrupted after about 12 Gyr.*

(see also [23]). At most five percent of its initial mass is still in a bound core. However the spread in the number of disc crossings is relatively wide and a significant fraction of microhalos should still have surviving cores. Mass loss is nevertheless important: microhalos maintaining more than 50% of their initial mass should be rare. Figure 1 illustrates the mass loss, where the solid curve shows the disruption of a typical microhalo with 80 disk crossings in 10 Gyr at the radius of the sun.

However, disk crossing is not the only source of dynamical disruption. While orbiting the galaxy, a microhalo is under the constant influence of the global Galactic potential, and tidal forces will act so that the microhalo's structure becomes elongated and unbound particles will form leading and trailing tidal streams. The detailed impact of tidal streaming depends on the orbit of the microhalo and on the shape of the host potential. In our simulations we use a disk potential that emerges from a density distribution of the form

$$\rho(R,z) \propto exp(-R/R_d)sech^2(z/z_d). \qquad (7)$$

Here R and z are the disk radius and the height respectively, which we set to be $R_d = 2.8$ kpc, $z_d = 0.4$ kpc. The disk mass is $M_d = 4.5 \cdot 10^{10} M_\odot$. In all our simulations the orbit of the microhalo is chosen to be roughly spherical with a distance of 7.9 kpc from the galactic center.

We cannot model both heating due to stellar interactions and tidal elongation at the same time since this would require following the motion of 50 billion disk stars. We therefore performed orbital simulations for three different cases: an initially completely undisturbed microhalo, a microhalo that first crossed the stellar field for 80 Myr and has lost about 60 percent of its mass, and a completely disrupted microhalo that spent more than 160 Myr in the stellar field.

The length of the tidal streams l due to the orbiting process can be crudely estimated with the relation $l(t) \sim \sigma_{mh} t$, where σ_{mh} is the velocity dispersion of the initial microhalo. For the initially unperturbed microhalo $\sigma_{mh} \sim 10^{-3}$ km/s, causing a stream length of roughly $l \sim 10$ pc after one Hubble time. For the initially completely disrupted microhalo $\sigma_{mh} \sim 10^{-2}$ km/s, and the stream length is about $l \sim 100$ pc after a Hubble time. These length scales agree well with our simulations (i.e. see Table II).

Orbiting in the galactic potential significantly reduces the mass of the microhalo (see dashed and dotted lines in Fig 1). However, the rate of tidal mass loss is suppressed as the tidal radius is steadily reduced. The central cusp of each dark matter microhalo has a very deep potential, as a consequence there is always a bound core remaining, even for a microhalo that has been heated in the stellar field before orbiting.

Comparing the curves in Fig 1 leads to the conclusion that disk crossing is the dominant disruption process and the only one that can lead to complete distruction of the microhalo. The step-like decrease of the curve is an indication of very close encounters that play a mayor role in the disruption process. Tidal stripping on the other hand can also significantly reduce the mass but it never completely distroys the microhalo because of its tightly bound inner core.

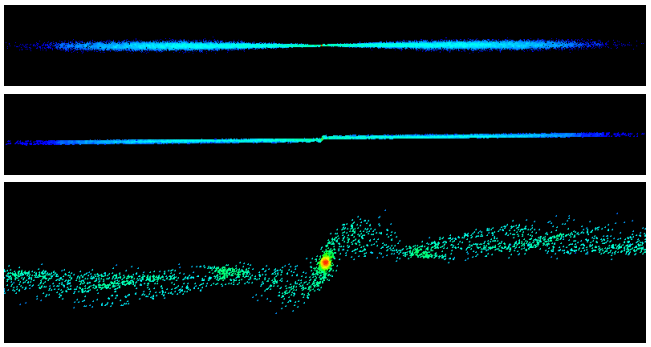

TABLE II: *Streaming microhalo after 10 Gyr on a roughly circular orbit around a Milky Way potential: The two images on the top show the sheet-like streams from the top and from the side (boxlength: 30 pc). The third image is a zoom in at the centre where the still bound core is visible (boxlength: 2 pc). For these pictures we have used an initially non disrupted microhalo with formation redshift 60.*

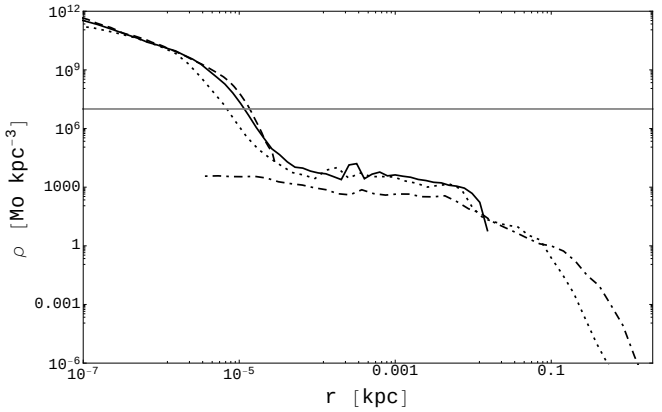

FIG. 2: *Stream densities of microhalos after an orbital time of ten Gyr. Before orbiting the microhalos have spent a time of 0 Myr (solid), 80 Myr (dotted) and 160 Myr (dashed-dotted) in the stellar field. There is still a visible bound core in the solid and the dotted profile but no more in the dashed-dotted one. The dashed line represents the density profile of a completely undisrupted microhalo that spend no time on an orbit. The horizontal gray line corresponds to the average dark matter density around the sun.*

Implications for Dark Matter Detection

In direct detection experiments the differential interaction rate is sensitive to the fine grained density and the velocity distribution of dark matter particles on A.U. scales [24, 25]. Substructures like microhalos can affect the interaction rate if they are abundant enough to have a substantial likelihood of existing in the solar neighbourhood and if their density is at least the same order of magnitude as the background dark matter density in this region, $\rho_{bg} \sim 10^7 M_\odot \mathrm{kpc}^{-3}$ (see for example [26]). Equally important, the phase space for the energy deposits associated with dark matter events will not be that appropriate for an isothermal halo if a single microhalo were to dominate the density distribution in the solar neighbourhood [13, 27].

Our results above suggest that none of these conditions are generally achieved. In Figure 2 we plot the stream densities of microhalos that crossed the stellar field for 0 Myr (solid line), 80 Myr (dotted line) and 160 Myr (dashed-dotted line), before orbiting in the galactic potential during 10 Gyr. The tidal streams of the initially unperturbed halo (solid line) have an average density of $\rho \sim 10^4 M_\odot \mathrm{kpc}^{-3}$, which is already negligibly low compared to the background. Only the very tiny core still maintains its initial density of $\rho \approx 10^{11} M_\odot \mathrm{kpc}^{-3}$. The initially disrupted microhalo (dashed-dotted line) has no more bound core. Its stream density is only at about $\rho \sim 10^2 - 10^3 M_\odot \mathrm{kpc}^{-3}$. The approach of first measuring the stellar disruption and then looking at tidal effects underestimates the stream density, since the microhalos get most of their heat energy right at the beginning. Therefore, the actual stream density of an average microhalo should be somewhere between the solid and the dashed-dotted line in Figure 2.

Since the stream densities are far below the value of the local galactic density, only a surviving core existing in the region of the earth would any effect upon direct detection. However, only about half of the microhalos still have bound cores because of disk crossing, and tidal effects further reduce the mass of the cores to less than ten percent of their original value. We note that any substructures orbiting primarily within the disk plane would be quickly destroyed by stellar encounters.

The chance of being in such an overdense region can then be optimistically estimated: An extrapolation of the subhalo mass function leads to a microhalo number density n_{mh} of about 500 pc^{-3} at the solar radius [7]. This number can be divided by two due to the disruption processes stated above and again by two since microhalos orbiting in the disk plane are completely disrupted. We then end up with the approximation of $n_{mh} \sim 100$ pc^{-3} at the solar radius. Each microhalo has a volume of about $V_{mh} \sim 10^{-9}$ pc^3 and therefore there is a chance of about 0.0001% of being in such an overdense region.

The streams of particles stripped from microhalos are coherent and long, thus it is appropriate to calculate their volume filling factor. Since the stream density is $\rho \sim 10^2 - 10^4 M_\odot \mathrm{kpc}^{-3}$ we expect that our solar system is criss-crossed with $f_b \times (10^3 - 10^5)$ streams, where $f_b \approx 0.1$ is the fraction of the local Galactic halo density that forms from substructures up to a solar mass. Larger substructures may be completely disrupted at the Sun's position in the Galaxy due to global disk shocking and tides [11]. The velocity dispersion within an average stream due to heating by disk stars is $\sigma \sim 10^{-2}$ km/s. Thus, the local density is determined by the superposition of a large number of independent streams, and the overall velocity distribution at the solar radius should be essentially Maxwellian, isotropic and smooth with no

spiky structure, as we would assume for a smooth halo model with no substructures. The signatures of streams could be only be detected experimentally with over several hundred events.

The case for indirect detection is somewhat different from that described above. In indirect detection experiments one tries to detect the annihilation products, such as gamma-rays, coming from the highest density dark matter regions, which is proportional to the square of the dark matter density times the volume of the region observed [28–31]. Consider a volume containing on average one microhalo $V \approx 10^{-2}$ pc^{-3}. The luminosity due to the smooth background is therefore $L_{bg} \propto V\rho^2 = 10^{-6} M_\odot^2 \mathrm{pc}^{-3}$, whereas the luminosity of a surviving microhalo core is

$$L_{mh} \propto V_{core}\rho_{core}^2 \approx 5 \times 10^{-7} M_\odot^2 pc^{-3}, \qquad (8)$$

where we have assumed a mean core density of $10^{10} M_\odot \mathrm{pc}^{-3}$. Thus the net boost factor due to microhalos is about 1.5, and stays below the detection limits of the FERMI experiment [32]. However this number is highly uncertain since it depends on extrapolations of both the substructure mass function and also on the microhalo internal density structure.

Finally, we consider the evolution of caustic sheets of particles within the Galactic halo. In the absence of fine grained heating, narrow sheets and folds will occupy regions of phase space within all collapsing CDM structures. As structures merge hierarchically, these caustic features become wrapped in phase space like a fine fabric that has been crumpled into a ball, the phase space density at any point being preserved. During the matter dominated epoch and before structure formation the velocity dispersion of WIMPs is given by

$$\sigma_\chi \sim 10^{-10} \left(\frac{100GeV}{m_\chi}\right)^{1/2} (1+z), \qquad (9)$$

where m_χ is the mass of the WIMP [12]. Since the first structures are collapsing at redshifts $z \sim 60$, we obtain a primordial velocity dispersion of $\sigma_\chi \sim 2$ cm/s. In the outer halo these features will persist, but in the vicinity of the sun heating by the disk stars will broaden the phase space and physical space distribution.

Since the energy increase is proportional to time (see equation (3)) we obtain a heating factor

$$\Delta\sigma = C\sqrt{t} \sim 10^{-3} \left(\frac{km}{s}\right) \sqrt{\frac{t}{Myr}}, \qquad (10)$$

where C has been determined via our simulations. For a microhalo on an average orbit we find an increase in the velocity dispersion of $\Delta\sigma \sim 12$ m/s which is much larger than the primordial velocity dispersion, effectively smearing out all caustic overdensities in the vicinity of the Galactic disk.

Conclusion

As the prospects for direct and indirect detection of WIMP dark matter improve with the development of new detectors, a renewed interest in the phase space distribution of dark matter particles has arisen. It has recently been shown that hierarchical clustering continues down to extremely small mass scales, so that most dark matter currently in the halo of our galaxy may have originated in microhalos with masses as small as $10^{-6} M_\odot$. Possibly dense surviving cores, tidal streams, and caustic structures might leave phase space sparsely populated, suggesting exciting new possibilities for novel signatures that differ from the traditional experimental assumption a smooth isothermal halo. However, our results imply that tidal effects and gravitational heating effectively wipe out any such signatures for Earth based detectors. Even though we find that a significant fraction of microhalos still have a bound core today, these overdense regions are too small to be relevant for detection experiments.

The disrupted material in the tidal streams is not dense enough to affect the detection signal. On Earth, there are about $7 \cdot 10^8$ dark matter particles per second streaming through our bodies (assuming $m_\chi \sim 100$ GeV), but they originate from over 10^3 streams coming from disrupted microhalos. The velocity dispersion in the streams is heated up through stellar interaction from initially 10^{-3} km/s to 10^{-2} km/s and therefore we expect an essentially smooth Maxwellian, with at most some spikes due to microhalos with an unusual orbital history [33, 34].

To summarise, our results imply that limits obtained on dark matter from detection experiments under the conservative assumption of a smooth halo with nearly Maxwellian density distribution remain valid, with a prefactor depending only on the average local dark matter density. The characteristic deviations from a Maxwellian distribution predicted from numerical simulations may be detected given sufficient detection statistics [35, 36]. Substructure cores still persist in the vicinity of the Sun to today, however the boost factor for indirect detection even for these is likely to be relatively small.

We thank Doug Potter for computational help. LMK acknowledges the hospitality of the ITP at the University of Zurich, where this work was initiated and the Pauli Center for Theoretical Studies for financial support. This research is supported by the Department of Energy Office of Science (Arizona), and the Swiss National Foundation.

[1] P. J. E. Peebles, Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations. *Astrophys. J.* **263**, L1-L5 (1982).

[2] B. Moore *et. al.*, Dark matter substructers in galactic

halos. *Astrophys J.* **524**, L19 (1999).
[3] S. Ghigna *et. al.*, Dark matter halos within clusters. *Mon. Not. R. Astron. Soc.* **300**, 146-162 (1998).
[4] A. M. Green, S. Hofmann, D. J. Schwarz, The power spectrum of SUSY-CDM on subgalactic scales. *Mon. Not. R. Astron. Soc.* **353**, L23-L27 (2004).
[5] L. Bergstrom, J. Edsjo, P. Gondolo, P. Ullio, Clumpy neutralino dark matter. *Phys. Rev. D* **59**, 043506 (1999).
[6] V. Berezinsky, V. Dokuchaev, Y. Eroshenko, Small-scale clumps in the galactic halo and dark matter annihilation. *Phys. Rev. D* **68**, 103003 (2003).
[7] J. Diemand, B. Moore, J. Stadel, Earth-mass dark-matter haloes as the first structures in the early Universe. *Nature* **433**, 389-391 (2005).
[8] S. M. Koushiappas, The detection of subsolar mass dark matter halos. *New J. Phys.* **11**, 105012 (2009).
[9] S. Hofmann, D. J. Schwarz, H. Stöcker, Damping scales of neutralino cold dark matter. *Phys. Rev. D* **64**, 083507 (2001).
[10] E. Bertschinger, Effects of cold dark matter decoupling and pair annihilation on cosmological perturbations. *Phys. Rev. D* **74**, 063509 (2006).
[11] E. D'Onghia, V. Springel, L. Hernquis, D. Keres, Substructure depletion in the Milky Way halo by the disk. *Astrophys. J.* **709**, 1138 (2010).
[12] P. Sikivie, The caustic ring singularity. *Phys.Rev. D* **60** 063501 (1999).
[13] C. J. Copi, L. M. Krauss, D. Simmons-Duffin, S. R. Stroiney, Assessing alternatives for directional detection of a WIMP halo. *Phys. Rev. D* **75**, 023514 (2007).
[14] The CDMS Collaboration: Z. Ahmed, *et. al.*, Results from the final exposure of the CDMS II experiment. arXiv:0912.3592v1 [astro-ph.CO] (2009).
[15] R. Bernabei *et. al.*, Results from the DAMA/LIBRA experiment. *J. of Physics* **203**, 012003 (2010).
[16] H. Zhao, D. Hooper, G. W. Angus, J. E. Taylor, J. Silk, Tidal disruption of the first dark microhalos. *Astrophys. J.* **654**, 697-701 (2007).
[17] T. Goerdt, O.Y. Gnedin, B. Moore, J. Diemand, J. Stadel, The survival and disruption of cold dark matter microhaloes: implications for direct and indirect detection experiments. *Mon. Not. R. Astron. Soc.* **375**, 191-198 (2006).
[18] B. Moore, N. Katz, G. Lake, A. Dressler, A. Oemler, Galaxy harassment and the evolution of clusters of galaxies. *Nature* **379**, 613 (1996).
[19] J. Binney, S. Tremaine, *Galactic Dynamics.* (Princeton University Press, Princeton, 2. ed. 2008).
[20] L. Spitzer, *Dynamical Evolution of Globular Clusters.* (Princeton University Press, Princeton, 1987).
[21] M. Zemp, B. Moore, J. Stadel, C. M. Carollo, P. Madau, Multimass spherical structure models for N-body simulations. *Mon. Not. R. Astron. Soc.* **386**, 1543-1556 (2008).
[22] L. M. Widrow, J. Dubinski, Equilibrium disk-bulge-halo models for the Milky Way and Andromeda galaxies. *Astrophys. J.* **636**, 838 (2005).
[23] A. M. Green, S. P. Goodwin, On mini-halo encounters with stars. *Mon. Not. R. Astron. Soc.* **375**, 1111-1120 (2007).
[24] G. Jungman, M. Kamionkowski, and K. Griest, Supersymmetric dark matter. *Phys. Rept.* **267**, 195 (1996)
[25] C. J. Copi, J. Heo, & L. M. Krauss, Directional sensitivity, WIMP detection, and the galactic halo, *Physics Letters* **B461**, 43 (1999)
[26] R. Catena, P. Ullio, A novel determination of the local dark matter density. arXiv:0907.0018 [astro-ph.CO] (2009).
[27] C. J. Copi, & L. M. Krauss, Angular signatures for galactic halo weakly interacting massive particle scattering in direct detectors: Prospects and challenges, *Phys. Rev. D*, **63**, 043507 (2001)
[28] G. Lake, Detectability of gamma-rays from clumps of dark matter. *Nature*, **346**, 39-40 (1990).
[29] C. Calcaneo-Roldan, B. Moore, Surface brightness of dark matter: Unique signatures of neutralino annihilation in the galactic halo. *Phys. Rev. D* **62**, 123005 (2000).
[30] FERMI-LAT Collaboration: A. A. Abdo *et. al.*, Constraints on cosmological dark matter annihilation from the Fermi-LAT isotropic diffuse gamma-ray measurement. arXiv:1002.4415v1 [astro-ph.CO] (2010).
[31] G. Bertone, D. Hooper, J. Silk, Particle dark matter: evidence, candidates and constraints *Phys. Rept.* **405**, 279-390 (2005).
[32] L. Pieri, E. Branchini, S. Hofmann, Difficulty of detecting minihalos via rays from dark matter annihilation. *Phys. Rev. Lett.* **95**, 211301 (2005).
[33] M. Kuhlen *et. al.*, Dark matter direct detection with non-Maxwellian velocity structure. *JCAP* **02**, 030 (2010).
[34] M. Vogelsberger, S. D. M. White, Streams and caustics: the fine-grained structure of LCDM haloes. arXiv:1002.3162v1 [astro-ph.CO] (2010).
[35] S. Hansen, B. Moore, M. Zemp, Stadel, J. A universal velocity distribution of relaxed collisionless structures. *JCAP*, **01**, 014H (2006).
[36] J.D. Vergados, S.H. Hansen, O. Host, Impact of going beyond the Maxwell distribution in direct dark matter detection rates. *Phys.Rev. D* **77**, 023509 (2008).

4

Paper II: WDM structure formation

In this second part of the thesis we are exploring the non linear structure formation within a ΛWDM universe. Dark matter particles are called warm, if they have small but non negligible velocity dispersion. From an astrophysical point of view such a dark matter candidate is promising, since it changes the small scale picture of the Universe, where CDM might have inconsistencies. The large scale behavior on the other hand stays unchanged and reproduces the stringent CMB and galaxy survey measurements.

The following manuscript compares the WDM halo model with numerical simulations. It analyzes the halo mass function, the bias, and the concentrations in different WDM scenarios and implements the findings into the halo model. The resulting prediction for the nonlinear power spectrum has a reduced error of about 5% and is now competitive with the fitting function of Viel et al. [40].

The paper has been published 2012 in the Monthly Notices of the Royal Astronomical Society [41].

Mon. Not. R. Astron. Soc. **000**, 1–?? (2011) Printed 4 May 2012 (MN LATEX style file v2.2)

Nonlinear evolution of cosmological structures in Warm Dark Matter models

Aurel Schneider[1⋆], Robert E. Smith[1,2], Andrea V. Macciò[3], and Ben Moore[1]
[1] *Institute for Theoretical Physics, University of Zurich, Zurich CH 8057*
[2] *Argelander-Institute for Astronomy, Auf dem Hügel 71, D-53121 Bonn, Germany*
[3] *Max Planck Institut für Astronomie, Königsstuhl 17, D-69117 Heidelberg, Germany*

4 May 2012

ABSTRACT
The dark energy dominated warm dark matter (WDM) model is a promising alternative cosmological scenario. We explore large-scale structure formation in this paradigm. We do this in two different ways: with the halo model approach and with the help of an ensemble of high resolution N-body simulations. Combining these quasi-independent approaches, leads to a physical understanding of the important processes which shape the formation of structures. We take a detailed look at the halo mass function, the concentrations and the linear halo bias of WDM. In all cases we find interesting deviations with respect to CDM. In particular, the concentration-mass relation displays a turnover for group scale dark matter haloes, for the case of WDM particles with masses of the order $m_{\rm WDM} \sim 0.25$keV. This may be interpreted as a hint for top-down structure formation on small scales. We implement our results into the halo model and find much better agreement with simulations. On small scales the WDM halo model now performs as well as its CDM counterpart.

Key words: Cosmology: theory - large-scale structure of the Universe - dark matter

1 INTRODUCTION

Over the last decade the vacuum energy dominated cold dark matter (hereafter ΛCDM) scenario, has emerged as a standard model for cosmology. This owes largely to the combination of information from galaxy clustering surveys such as the 2dFGRS and SDSS with WMAP measurements of the temperature anisotropies in the microwave background (Cole et al. 2005; Tegmark et al. 2006; Komatsu et al. 2011). However, the nature of the two dark components in the ΛCDM model are still completely unknown and it is therefore important to keep exploring alternative models and test their compatibility with observations.

In the ΛCDM model the dark matter is assumed to be composed of heavy, cold thermal relic particles that decoupled from normal matter very early in the history of the Universe (Peebles 1982; Blumenthal et al. 1984; Kolb & Turner 1990; Jungman et al. 1996). Whilst there is a large body of indirect astrophysical evidence that strongly supports CDM, there are some hints that it has shortcomings. Firstly, CDM galaxy haloes contain a huge number of subhaloes (Moore et al. 1999; Diemand & Kuhlen 2008; Springel et al. 2008; Stadel et al. 2009), while observations indicate that only relatively few satellite galaxies exist around the Milky Way and M31 (Moore et al. 1999; Klypin et al. 1999). Secondly, the highest resolution halo simulations show that the slope of the inner density profile decreases linearly at smaller radii (Navarro et al. 1997; Moore et al. 1999; Diemand et al. 2004; Springel et al. 2008; Stadel et al. 2009), whereas the density profiles inferred from galaxy rotation curves are significantly shallower (Moore et al. 1999) (and for recent studies see Swaters et al. 2003; Salucci et al. 2007; de Blok et al. 2008; Gentile et al. 2009, and references there in). Thirdly, the observed number of dwarf galaxies in the voids appears to be far smaller than expected from CDM (Peebles 2001; Tikhonov et al. 2009; Peebles & Nusser 2010). Another example is the excess in the prediction of dwarf galaxy concentrations (Lovell et al. 2011). Whilst, it has become clear that some of these discrepancies might be resolved through an improved understanding of galaxy formation, they have led some to consider changes to the ΛCDM paradigm.

One possible solution might be warm dark matter (WDM) (Bond & Szalay 1983; Bardeen et al. 1986; Bode et al. 2001). In this scenario, the dark particle is considered to be lighter than its CDM counterpart, and so remains relativistic longer and also retains a thermal velocity. Since WDM particles are collisionless and decouple early, they may 'free-stream' or diffuse out of perturbations whose

⋆ Email: aurel@physik.uzh.ch

size is smaller than the Jeans' length[1] in the radiation dominated Universe (Kolb & Turner 1990). This free-streaming of the WDM particles acts to damp structure formation on small scales. Two potential candidates are the sterile neutrino (Dodelson & Widrow 1994; Shaposhnikov & Tkachev 2006), and the gravitino (Ellis et al. 1984; Moroi et al. 1993; Kawasaki et al. 1997; Gorbunov et al. 2008), both of which require extensions of the standard model of particle physics.

Recent observational constraints have suggested that sterile neutrinos can not be the dark matter: the Lyman alpha forest (Seljak et al. 2006; Boyarsky et al. 2009a) and QSO lensing (Miranda & Macciò 2007) bounds are $m_{\nu_s} > 8$keV, whilst those from the X-ray background are $m_{\nu_s} < 4$keV (Boyarsky et al. 2008)[2]. However, a more recent assessment has suggested that a better motivated particle physics model based on resonant production of the sterile neutrino, may evade these constraints: the Lyman alpha forest bound is brought down to $m_{\nu_s} \gtrsim 2$keV and the X-ray background is pushed to $m_{\nu_s} < 50$keV (for very low mixing angles) (Boyarsky et al. 2009b). It therefore seems that additional, independent methods for constraining the ΛWDM scenario would be valuable.

In Markovic et al. (2010) and Smith & Markovic (2011), it was proposed that the ΛWDM scenario could be tested through weak lensing by large-scale structure. The advantage of such a probe is that it is only sensitive to the total mass distribution projected along the line of sight. However, to obtain constraints on the WDM particle mass, an accurate model for the nonlinear matter clustering is required. In these papers, an approach based on the halo model was developed. Accurate predictions from this model require: detailed knowledge of the abundance of dark matter haloes, their spatial large-scale bias, and their density profiles. In these studies, it was assumed that the semi-analytic methods, which were developed for CDM, would also apply to WDM.

In this paper we perform a series of very high resolution CDM and WDM N-body simulations with the specific aim of exploring the halo model ingredients in the ΛWDM scenario. Over the past decade, there have been a limited number of numerical simulation studies of nonlinear structure formation in the WDM model (Colombi et al. 1996; Moore et al. 1999; Colín et al. 2000; White & Croft 2000; Avila-Reese et al. 2001; Bode et al. 2001; Bullock et al. 2002; Zentner & Bullock 2003; Colín et al. 2008; Zavala et al. 2009; Macciò & Fontanot 2010; Lovell et al. 2011; Viel et al. 2011; Dunstan et al. 2011). In most of these previous studies, conclusions have been drawn from object-by-object comparison of a relatively small number of haloes simulated in boxes of typical size $L = 25\,h^{-1}$Mpc. In this work we are more interested in the overall impact that the WDM hypothesis has on the statistical properties of large-scale structures. We therefore simulate boxes that are 10 times larger than have been typically studied before, hence having roughly ~ 1000 times larger sampling volume. This means, that our conclusions will have greater statistical weight, than those from previous studies. Furthermore, our results should be less susceptible to finite volume effects, which can lead to underestimates of the nonlinear growth.

The paper is structured as follows: In §2 we provide a brief overview of the salient features of linear theory structure formation in the WDM model and we review the halo model approach. In §3 we describe the N-body simulations. In §4 we explore the main ingredients of the halo model: the halo mass function, bias and density profiles. In §5 we compare the halo model predictions for the matter power with our measurements from the simulations. Finally, in §6 we summarize our findings.

2 THEORETICAL BACKGROUND

In this section we summarize the linear theory for WDM and the nonlinear halo model in this framework.

2.1 Linear theory evolution of WDM

The physics of the free-streaming or diffusion of collisionless particles out of dark matter perturbations has been discussed in detail by Bond & Szalay (1983)[3]. An estimate for the free-streaming length can be obtained, by computing the comoving length scale that a particle may travel up until matter-radiation equality ($t_{\rm EQ}$). At this point, the Jeans' length drops dramatically and perturbations may collapse under gravity. A simple formula for this is given by Kolb & Turner (1990):

$$\lambda_{\rm fs} = \int_0^{t_{\rm EQ}} \frac{v(t)dt}{a(t)} \approx \int_0^{t_{\rm NR}} \frac{cdt}{a(t)} + \int_{t_{\rm NR}}^{t_{\rm EQ}} \frac{v(t)dt}{a(t)} \ , \qquad (1)$$

where $t_{\rm NR}$ is the epoch when the WDM particles become non-relativistic, which occurs when $T_{\rm WDM} < m_{\rm WDM}c^2/3k_{\rm B}$, where $T_{\rm WDM}$ and $m_{\rm WDM}$ are the characteristic temperature and mass of the WDM particles. In the relativistic case, the mean peculiar velocity of the particle is simply $v(t) \sim c$. In the non-relativistic regime its momentum simply redshifts with the expansion: $v \propto a(t)^{-1}$. This leads to:

$$\lambda_{\rm fs} \approx r_{\rm H}(t_{\rm NR}) \left[1 + \frac{1}{2}\log\frac{t_{\rm EQ}}{t_{\rm NR}}\right] \ , \qquad (2)$$

where $r_{\rm H}(t_{\rm NR})$ is the comoving size of the horizon at $t_{\rm NR}$. On inserting typical values for $t_{\rm NR}$ we find the scaling:

$$\lambda_{\rm fs} \approx 0.4 \left(\frac{m_{\rm WDM}}{\rm keV}\right)^{-4/3} \left(\frac{\Omega_{\rm WDM}h^2}{0.135}\right)^{1/3} \left[h^{-1}{\rm Mpc}\right] \qquad (3)$$

However, the real situation is more complex than this, since fluctuations inside the horizon grow logarithmically during radiation domination via the Meszaros effect and free-streaming does not switch off immediately after $t_{\rm EQ}$. To understand the collisionless damping in more detail, one must numerically solve the coupled Einstein-Boltzmann system of equations for the various species of matter and radiation.

[1] Although originally defined in the context of gas dynamics, the Jeans length can be generalized to collisionless systems by replacing the sound speed with the velocity dispersion. The reason for this tight analogy lies in the linearized equation of perturbations, which has the same structure for gas and collisionless fluids (see Peebles 1982, for more details).

[2] Lower bounds on the mass of a fully thermalized WDM particle can be obtained using Eq. (6) (see Viel et al. 2005).

[3] For some recent theoretical treatments of WDM, also see Boyanovsky (2010) and de Vega & Sanchez (2010, 2011)

Several fitting formulae for the WDM density transfer function have been proposed (Bardeen et al. 1986; Bode et al. 2001) and here we adopt the formula in Viel et al. (2005):

$$T_{\rm WDM}(k) = \left[\frac{P_{\rm lin}^{\rm WDM}}{P_{\rm lin}^{\rm CDM}}\right]^{1/2} = \left[1+(\alpha k)^{2\mu}\right]^{-5/\mu}, \quad (4)$$

with $\mu = 1.12$ as well as

$$\alpha = 0.049\left[\frac{m_{\rm WDM}}{\rm keV}\right]^{-1.11}\left[\frac{\Omega_{\rm WDM}}{0.25}\right]^{0.11}\left[\frac{h}{0.7}\right]^{1.22}{\rm Mpc/h}. \quad (5)$$

Note that in the above we are assuming that the WDM particle is fully thermalized. Following Viel et al. (2005), the masses of sterile neutrino WDM particles m_{ν_s} can be obtained from $m_{\rm WDM}$ through the formula:

$$m_{\nu_s} = 4.43{\rm keV}\left(\frac{m_{\rm WDM}}{1{\rm keV}}\right)^{4/3}\left(\frac{w_{\rm WDM}}{0.1225}\right)^{-1/3}. \quad (6)$$

The characteristic length-scale α is related to the free-streaming scale $\lambda_{\rm fs}$, and we shall therefore make the definition that $\alpha \equiv \lambda_{\rm fs}^{\rm eff}$ is an *effective* free-streaming length scale. The length-scale $\lambda_{\rm fs}^{\rm eff}$ can be used to introduce the 'free-streaming' mass scale:

$$M_{\rm fs} = \frac{4\pi}{3}\bar{\rho}\left(\frac{\lambda_{\rm fs}^{\rm eff}}{2}\right)^3, \quad (7)$$

where $\bar{\rho}$ is the background density of the universe. This mass scale is important as it defines the scale below which initial density perturbations are completely erased.

We can define yet another length scale: the 'half-mode' length scale $\lambda_{\rm hm}$. This corresponds to the length scale at which the amplitude of the WDM transfer function is reduced to 1/2. From Eq. (4) we find:

$$\lambda_{\rm hm} = 2\pi\lambda_{\rm fs}^{\rm eff}\left(2^{\mu/5}-1\right)^{-1/2\mu} \approx 13.93\lambda_{\rm fs}^{\rm eff}. \quad (8)$$

This length scale leads us to introduce another mass scale, the half-mode mass scale:

$$M_{\rm hm} = \frac{4\pi}{3}\bar{\rho}\left(\frac{\lambda_{\rm hm}}{2}\right)^3 \approx 2.7\times10^3 M_{\rm fs}. \quad (9)$$

This mass scale is where we expect the WDM to first affect the properties of dark matter haloes. In passing, this partly explains the claims made by Smith & Markovic (2011), that, for instance, the mass function of haloes would be significantly suppressed on mass scales $M \sim 100 M_{\rm fs}$.

In Fig. 1 we show the relation between $M_{\rm fs}$, $M_{\rm hm}$ and the mass of the WDM particle candidate for our adopted cosmological model. Three cases of relevance are apparent: $M > M_{\rm hm}$, and haloes form hierarchically through accreting material; $M_{\rm hm} > M > M_{\rm fs}$ and for these haloes the hierarchy may fail with low mass haloes forming at the same time as higher mass haloes; finally $M_{\rm fs} > M$ no halo formation, unless through the fragmentation of larger structures. While the growth of overdensities is not affected above $M_{\rm hm}$, it is suppressed between $M_{\rm fs}$ and $M_{\rm hm}$, and should simply not take place below $M_{\rm fs}$.

2.2 Nonlinear evolution of WDM: the halo model

Cosmological structure formation is a very complicated, highly nonlinear process that requires numerical simulation for a full understanding. However, the halo model approach

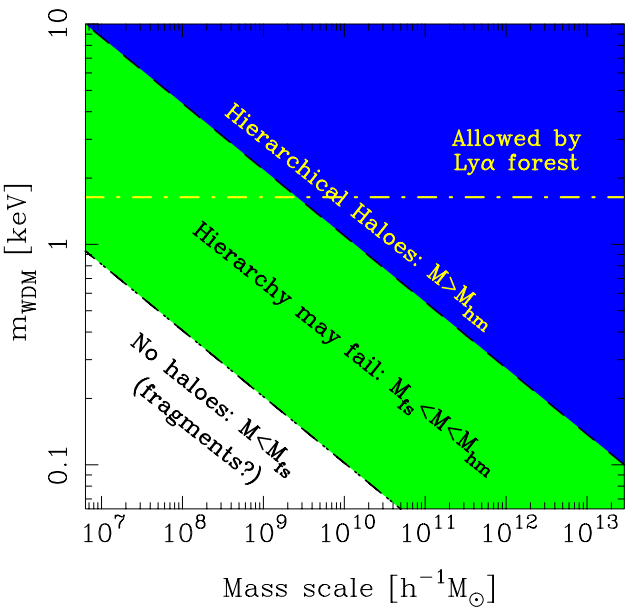

Figure 1. Free-streaming mass-scale ($M_{\rm fs}$) and half-mode mass scale ($M_{\rm hm}$) as a function of the mass of the WDM particle ($m_{\rm WDM}$). Haloes with masses $M > M_{\rm hm}$, may form hierarchically (upper right solid blue region). For haloes with masses $M_{\rm hm} > M > M_{\rm fs}$, hierarchical structure growth may fail (middle green region). For haloes with masses $M < M_{\rm fs}$, these may not form hierarchically since their initial peaks are completely erased (lower right empty region). However it is possible that such objects may emerge through fragmentation. The yellow dot-dash line denotes the current $m_{\rm WDM}$ allowed by the Lyman alpha forest (Boyarsky et al. 2009a) (note that we have rescaled $m_{\nu_s} \rightarrow m_{\rm WDM}$ using Eq. (6)).

gives a simplified analytical description of structure formation, which leads to surprisingly good results (Cooray & Sheth 2002, and references therein). Recently, the halo model has been adapted for the WDM cosmological model by Smith & Markovic (2011) and we now summarize their basic approach.

The main idea of the halo model in WDM is to separate the density field into a halo component, adding up all bound structure, and a smooth component, standing for all matter, that has not collapsed due to free streaming. This is different to the standard approach of the CDM halo model, where all matter is supposed to be in bound structures.

Thus the WDM density field has the form,

$$\rho(\mathbf{x}) = \rho_s(\mathbf{x}) + \sum_{i=1}^{N} M_i u(|x-x_i|, M_i), \quad (10)$$

where ρ_s is the smooth part of the density field and $u(x, M) = \rho_h(x|M)/M$ is the mass normalized density profile. The average densities of the smooth and the bound components are then given by

$$\langle\rho\rangle = \bar{\rho} = \bar{\rho}_{\rm s} + \bar{\rho}_{\rm h}, \qquad \bar{\rho}_{\rm h} = f\bar{\rho}, \quad (11)$$

where f is the fraction of matter in bound objects. This can be calculated by integrating over the halo mass function weighted by halo mass:

$$f = \frac{1}{\bar{\rho}}\int_0^\infty d\log M M\frac{dn}{d\log M}, \quad (12)$$

where $dn = n(M)dM$ is the abundance of WDM haloes of mass M in the interval dM. The fraction f is equal to unity in a perfectly hierarchical universe and drops below unity as soon as the mass function is suppressed due to the free streaming. In a WDM universe the amount of suppression depends on the mass of the WDM particle.

The power spectrum $P(k)$ is defined by the relation

$$\langle \delta(\mathbf{k})\delta(\mathbf{k}')\rangle \equiv (2\pi)^3 \delta_D(\mathbf{k}+\mathbf{k}')P(k) \ , \quad (13)$$

where δ_D is the three dimensional Dirac delta function and $\delta(k)$ is the Fourier transform of the matter overdensity $\delta(\mathbf{x}) \equiv (\rho(\mathbf{x}) - \overline{\rho})/\overline{\rho}$. In terms of the different density components, we can write:

$$\delta(k) = f\delta_{\rm h}(k) + (1-f)\delta_{\rm s}(k), \quad (14)$$

where $\delta_\chi \equiv (\rho_\chi - \overline{\rho}_\chi)/\overline{\rho}_\chi$ with $\chi \in \{{\rm h}, {\rm s}\}$. The power spectrum of the halo model can now be determined by adding up the power spectra of the different density components as well as their cross terms, giving

$$P(k) = (1-f)^2 P_{\rm ss}(k) + 2(1-f)fP_{\rm sh}(k) + f^2 P_{\rm hh}(k) \ . \quad (15)$$

The term $P_{\rm hh}$ represents the power spectrum of matter trapped in haloes, the term $P_{\rm ss}$ designates the power spectrum of the smooth component and the term $P_{\rm sh}$ denotes the cross-power spectrum between haloes and the smooth field.

The term $P_{\rm hh}$ can be separated into one- and two-halo terms, which describe the power coming from the same halo, and the one coming from distinct haloes, respectively. It can be expressed as:

$$P_{\rm hh}(k) = P^{\rm 2h}_{\rm hh}(k) + P^{\rm 1h}_{\rm hh}(k) \ ; \quad (16)$$

$$P^{\rm 2h}_{\rm hh}(k) = \prod_{i=1}^{2} \left\{ \int_0^\infty \frac{dM_i}{\bar{\rho}_{\rm h}} M_i n(M_i) u(M_i) \right\} \times P^{\rm c}_{\rm hh}(k|M_1, M_2), \quad (17)$$

$$P^{\rm 1h}_{\rm hh}(k) = \frac{1}{\bar{\rho}^2_{\rm h}} \int_0^\infty dM n(M) M^2 u^2(k|M) \ , \quad (18)$$

where $u(k|M)$ is the Fourier transform of the mass normalized density profile. In Eq. (17) we have introduced the power spectrum of halo centers $P^{\rm hh}_c(k|M_1, M_2)$, which in general is a complicated function of k and the halo masses M_1 and M_2. However, if we neglect halo exclusion and assume linear biasing with respect to the linear mass density, then we may write this as,

$$P^c_{\rm hh}(k|M_1, M_2) \sim b_1(M_1)b_1(M_2)P_{\rm lin}(k) \ . \quad (19)$$

In this case, the function is separable and this considerably simplifies the integrals in Eq. (17). This approximation breaks down on small, nonlinear scales, but on these scales, the two-halo term is sub-dominant. The error induced by this approximation (19) is most apparent at quasi-linear scales ($k \sim [0.1, 1.0]\, h\,{\rm Mpc}^{-1}$) and is $\lesssim 30\%$. It is possible to lower this error to $\lesssim 5\%$ by using higher order perturbation theory techniques and by including halo exclusion (see for example Smith et al. 2011). An easy but not fully consistent way of reducing the error down to $\lesssim 10\%$ is to do the following replacement in Equation (19):

$$P_{\rm lin}(k) \rightarrow P_{\rm halofit}(k) W_{\rm TH}(kR), \qquad R \simeq 2\, h^{-1}{\rm Mpc}, \quad (20)$$

where $W_{\rm TH}$ is the window function defined in §4.1 and

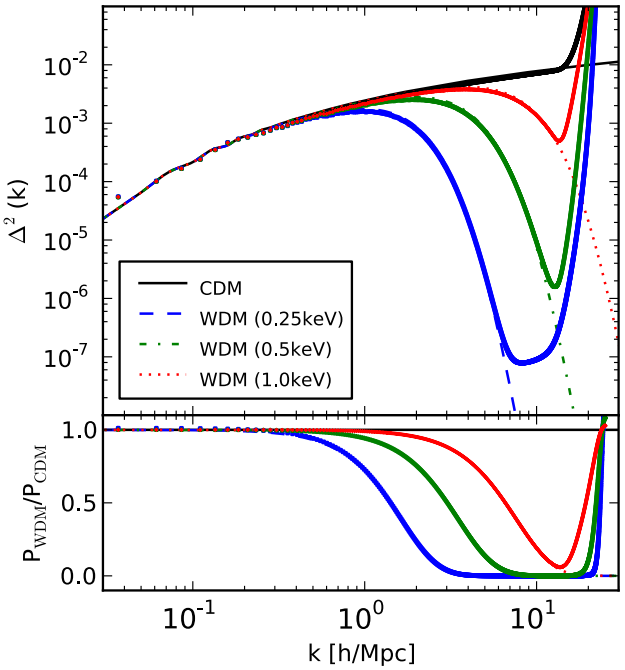

Figure 2. Linear power spectra as a function of wavenumber in the CDM and WDM scenarios, at the initial redshift ($z = 49$) of the simulations. *Top panel*: absolute dimensionless power: $\Delta^2 = k^3P(k)/2\pi^2$. The lines denote the linear power spectrum where $m_{\rm WDM} \in \{\infty, 0.25,\ 0.5,\ 1.0\}$ keV. Points denote the power spectra measured from the initial conditions of the $N = 1024^3$ simulations. *Bottom panel*: ratio of the initial WDM and CDM power spectra. Points and lines unchanged.

$P_{\rm halofit}$ is the power spectrum calculated by the `halofit` code (Smith et al. 2003).

The halo-smooth power spectrum is given by:

$$P_{\rm sh}(k) = \frac{1}{\bar{\rho}_{\rm h}} \int dM n(M) M u(k|M) P^c_{\rm hs}(k|M) \ , \quad (21)$$

where $P^c_{\rm hs}(k|M)$ is the power spectrum of the halo centers with respect to the smooth mass field. On assuming that the smooth field and the halo density field are linearly biased with respect to the linear density field, we are lead to the relation:

$$P^c_{\rm sh}(k|M) \sim b_s b(M) P_{\rm lin}(k) \ , \quad (22)$$

where b_s is the linear bias of the smooth matter field defined in §4.3. Finally, the smooth field auto-power spectrum is given by

$$P_{\rm ss}(k) = b_{\rm s}^2 P_{\rm lin}(k) \ . \quad (23)$$

In order to reduce the error we can again replace the linear power spectrum in the Equations (22) and (23), following the recipe of relation (20).

On combining these power spectra, weighted by the correct functions of their mass fractions, à la Eq. (15), we find the total halo model prediction for the nonlinear matter power spectrum in the WDM model.

Sim label	$m_{\rm WDM}$ [keV]	$M_{\rm fs}$ $[h^{-1}M_\odot]$	$M_{\rm hm}[h^{-1}M_\odot]$	L $[h^{-1}{\rm Mpc}]$	$N_{\rm part}$	m_p $[h^{-1}M_\odot]$	$l_{\rm soft}[h^{-1}{\rm kpc}]$
CDM-S					256^3	7.57×10^{10}	20
CDM-M	∞	0	0	256	512^3	9.45×10^{9}	10
CDM-L					1024^3	1.18×10^{9}	5
WDM-1.25-S					256^3	7.57×10^{10}	20
WDM-1.25-M	1.25	2.3×10^{6}	6.3×10^{9}	256	512^3	9.45×10^{9}	10
WDM-1.25-L					1024^3	1.18×10^{9}	5
WDM-1.0-S					256^3	7.57×10^{10}	20
WDM-1.0-M	1.0	4.9×10^{6}	1.3×10^{10}	256	512^3	9.45×10^{9}	10
WDM-1.0-L					1024^3	1.18×10^{9}	5
WDM-0.75-S					256^3	7.57×10^{10}	20
WDM-0.75-M	0.75	1.3×10^{7}	3.4×10^{10}	256	512^3	9.45×10^{9}	10
WDM-0.75-L					1024^3	1.18×10^{9}	5
WDM-0.5-S					256^3	7.57×10^{10}	20
WDM-0.5-M	0.5	4.9×10^{7}	1.3×10^{11}	256	512^3	9.45×10^{9}	10
WDM-0.5-L					1024^3	1.18×10^{9}	5
WDM-0.25-S					256^3	7.57×10^{10}	20
WDM-0.25-M	0.25	5.0×10^{8}	1.3×10^{12}	256	512^3	9.45×10^{9}	10
WDM-0.25-L					1024^3	1.18×10^{9}	5

Table 1. WDM simulations. From left to right, columns are: simulation name (S=Small, M=Medium, L=Large); mass of WDM particle ($m_{\rm WDM}$); free-streaming mass-scale ($M_{\rm fs}$); half-mode mass-scale ($M_{\rm hm}$); simulation box-size (L); number of particles ($N_{\rm part}$); mass of simulation particles ($m_{\rm p}$); comoving softening length ($l_{\rm soft}$).

3 N-BODY SIMULATIONS OF WDM

In order to study nonlinear structure growth in the WDM model, we have generated a suite of N-body simulations. These were executed on the `zBOX3` supercomputer at the University of Zürich. Each simulation was performed using `PKDGRAV`, a high order multipole tree-code with adaptive time stepping (Stadel 2001).

The cosmological parameters of the base ΛCDM model adopted, are consistent with the WMAP7 best-fit parameters (Komatsu et al. 2011) and we take: the energy-density parameters in matter, vacuum energy and baryons to be $\Omega_{\rm m}=0.2726$, $\Omega_\Lambda=0.7274$, $\Omega_{\rm b}=0.046$; the dimensionless Hubble parameter to be $h=0.704$; the primordial power spectral index and present day normalization of fluctuations to be $n_s=0.963$, and $\sigma_8=0.809$.

The CDM transfer function was generated using the code `CAMB` (Lewis et al. 2000). The linear power spectrum for each WDM model was then obtained by multiplying the linear CDM power spectrum by $T^2_{\rm WDM}(k)$ from Eq. (4). Initial conditions for each WDM model were then generated at redshift $z=49$ using the serial version of the publicly available `2LPT` code (Scoccimarro 1998; Crocce et al. 2006). In theory, we should also include a velocity dispersion due to the fact that the particles still retain a relic thermal velocity distribution. However, a quick calculation of the rms dispersion velocity, showed that these effects should be of marginal importance on scales $\gtrsim 50\,h^{-1}$kpc at the initial redshift, and of order $\gtrsim 1\,h^{-1}$kpc at the present day for $m_{\rm WDM}\geqslant 0.25\,[{\rm keV}]$. We therefore assume that their inclusion will be a second order effect and so at this stage we neglect them.

We generated initial conditions for a suite of simulations, one with a CDM particle and five with WDM particle masses $m_{\rm WDM}\in\{0.25,0.5,0.75,1.0,1.25\}$keV. For all runs, we set the box length $L=256\,h^{-1}$Mpc. This size is a compromise between choosing a box small enough to accurately capture small-scale structure formation and large enough to confidently follow the linear evolution of the box-scale modes. This makes it possible for us to check agreement with the linear theory and to measure linear bias.

Our simulations were also performed with three different mass resolutions: $N=\{256^3,\,512^3,\,1024^3\}$. This enables us to differentiate between genuine structures and spurious structures, which can collapse out of the initial particle lattice (cf. Wang & White 2007; Polisensky & Ricotti 2010). Full details of the suite of simulations are summarized in Table 1.

Dark matter haloes in the simulations were located using the Friends-of-Friends algorithm (Davis et al. 1985). We used a modified version of the `skid` code, with the linking length parameter set to the conventional value of $b=0.2$.

Fig. 2 compares the initial linear theory power spectra with the power spectra estimated from the initial conditions of the N-body simulations, for the case $N=1024^3$. These results show that, at the initial time, the WDM linear theory distribution of fluctuations has been correctly seeded. It also shows a spike in the measured power spectrum at $k=8\pi$ which is a consequence of the initial particle distribution on a grid.

Fig. 3 presents a pictorial view of the growth of structure in a selection of the simulations. The left column shows the density evolution in a slice through one of the CDM simulations. The central and right panels show the same but for the case of WDM particles with $m_{\rm WDM}=1.0\,{\rm keV}$ and $m_{\rm WDM}=0.25\,{\rm keV}$. From top to bottom the panels show results for $z=4.4$, 1.1 and 0.

4 HALO MODEL INGREDIENTS IN THE WDM SCENARIO

In this section we detail the halo model ingredients and show how they change in the presence of our benchmark set of WDM particle masses.

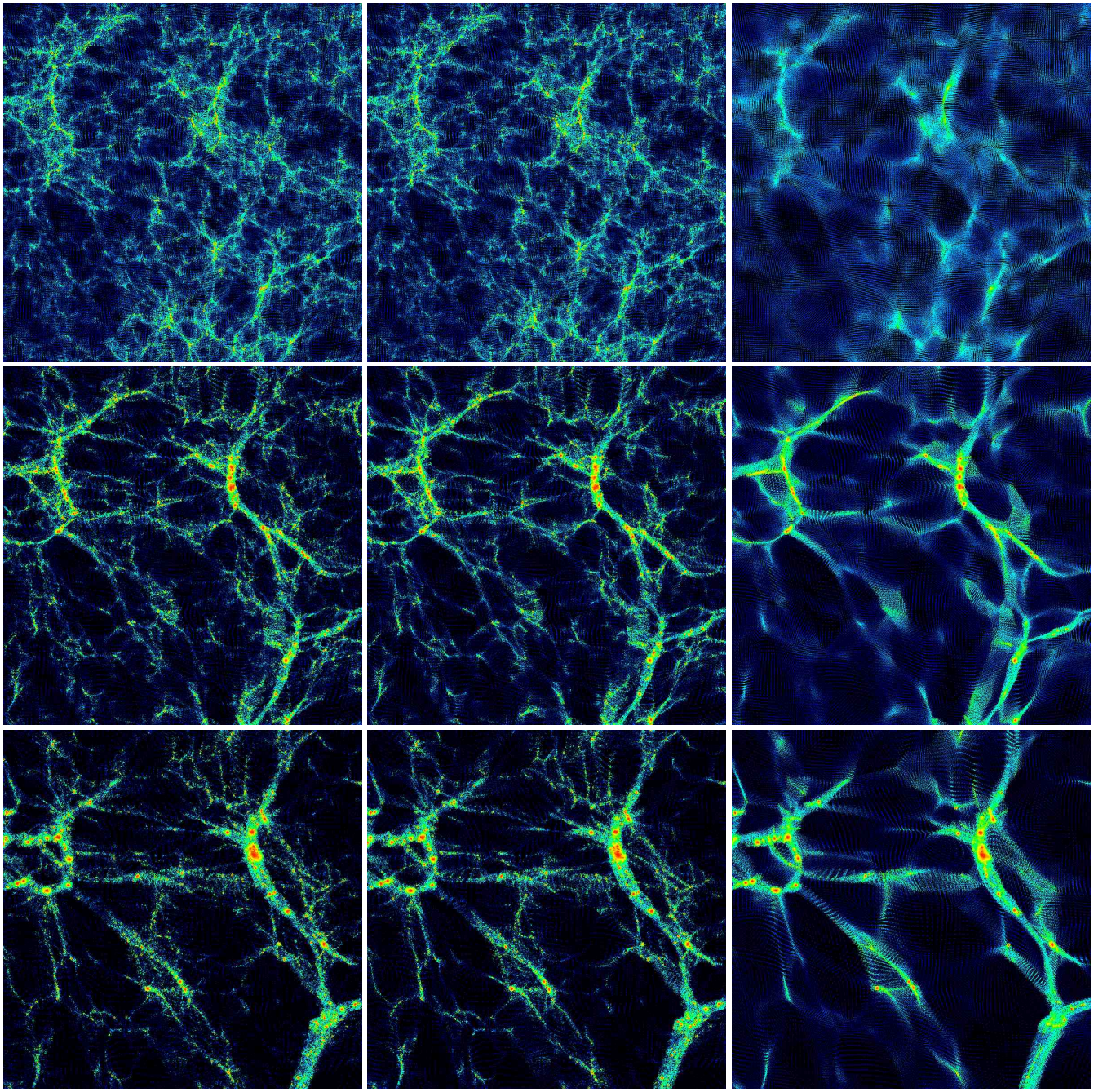

Figure 3. Density maps from the $N = 1024$ simulations with a length of $50\,h^{-1}$Mpc and a depth of $2.5\,h^{-1}$Mpc. From top to bottom: $z = 4.4$, $z = 1.1$ and $z = 0$. From left to right: CDM, WDM with $m_{\rm p} = 1.0$keV and WDM with $m_{\rm p} = 0.25$keV. Whilst the WDM effects are barely discernible in the middle panels, they are very prominent in the right panels, where the voids are noticeably emptier than in CDM.

4.1 Halo mass function

In CDM the halo mass function can be explored through the excursion set formalism (Press & Schechter 1974; Bond et al. 1991):

$$\frac{dn}{d\log M} = -\frac{1}{2}\frac{\overline{\rho}}{M} f(\nu) \frac{d\log\sigma^2}{d\log M}\ . \tag{24}$$

In the ellipsoidal collapse model of Sheth & Tormen (1999), $f(\nu)$ is given by

$$f(\nu) = A\sqrt{\frac{2q\nu}{\pi}}\left[1 + (q\nu)^{-p}\right] e^{-q\nu/2}, \quad \nu = \frac{\delta_c^2(t)}{\sigma^2(M)}, \tag{25}$$

with parameters: $p = 0.3$, $q = 0.707$ and with normalization parameter $A = 0.3222$. The linear theory collapse threshold is given by, $\delta_c(z) \equiv 1.686/D(z)$, where $D(z)$ is the linear theory growth function. The variance on mass scale M is,

$$\sigma^2(M) = \int \frac{{\rm d}^3\mathbf{k}}{(2\pi)^3} P_{\rm Lin}(k) W_{\rm TH}^2(kR)\ , \tag{26}$$

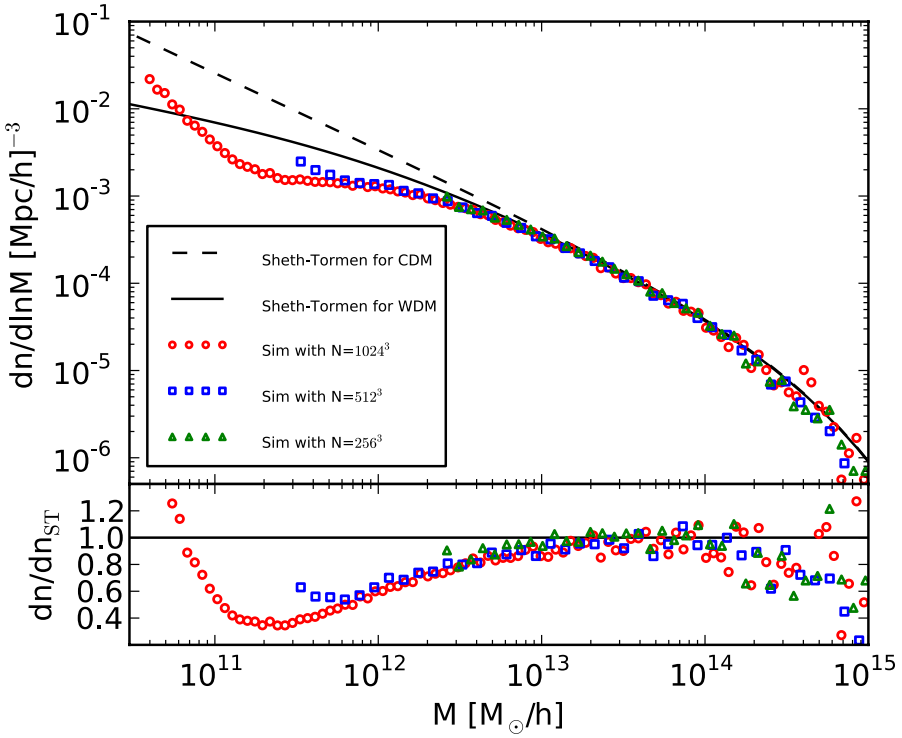

Figure 4. Measured mass function of the WDM simulations with $m_p = 0.25$ keV and three different resolutions. The measurements lie below the Sheth-Tormen prediction, a well known result that is discussed in section 4. The upturn of the mass function due to artificial haloes is visible in the simulations of high and medium resolution.

where $W_{\rm TH}(y) \equiv 3\left[\sin y - y\cos y\right]/y^3$ and where the mass scale and radius of the filter function are related through the relation: $M = 4\pi R^3\bar{\rho}/3$.

The main idea in the excursion set approach is that there is a monotonic mapping between the linear and nonlinear density perturbations, averaged over a randomly selected patch of points in the space. Further, the mapping can be calculated using the spherical or ellipsoidal collapse approaches. The density perturbation in the patch will collapse to form a virialized object when the linearly extrapolated density in the patch reaches a certain collapse threshold. Despite the fact that this approach does not trace the full complexity of nonlinear structure formation, the actual predictions are in close agreement with measurements from simulations. That is, at least for a CDM cosmology with well calibrated values for the ellipsoidal parameters p and q and a given halo finding algorithm. One important assumption, which is implicit within this framework, is that structure formation must proceed hierarchically.

In the WDM scenario, things are more complicated, since structure formation may not always proceed hierarchically. As described in §2, we can identify three regimes of interest: for $M > M_{\rm hm}$, the variance of WDM fluctuations becomes virtually indistinguishable from that for CDM, and the excursion set approach should be valid; for $M < M_{\rm fs}$ all primordial overdensities are erased through diffusion of particles during the epoch of radiation domination and we expect that no hierarchical halo formation will take place on these mass scales. In between, where $M_{\rm hm} > M > M_{\rm fs}$, the WDM overdensity field is suppressed, but there is still some power left that may enable hierarchical collapse to take place. It is not clear *a priori*, how the mass function behaves on these scales and whether the extended Press-Schechter approach remains valid. We now investigate this using our simulations.

In Fig. 4 we show the $z = 0$ mass function of dark matter haloes for the case of $m_{\rm WDM} = 0.25$keV. The figure demonstrates the behaviour of the mass function as the simulation resolution is increased from $N = 256^3$, to 512^3, to 1024^3 particles, denoted by the triangles, squares and circles, respectively. We can now see the effect of artificial clumping (cf. discussion in §3), which is manifest as the upturn of the curves at the low mass end of the mass function. One common approach to dealing with this artificial clumping is to assume that the simulations can be trusted down to the mass-scale just above the up-turn. We also find, in agreement with Wang & White (2007), that this mass-scale increases as $N^{1/3}$, i.e. the inter-particle spacing. In order to decrease the resolved mass by a factor of two, the particle resolution has to go up by a factor of eight. This is one of the main reasons why simulating WDM models is significantly more challenging than simulating CDM models.

Fig. 4 also shows the prediction of the halo mass function for CDM and for this WDM model, from the ST mass function. The figure clearly shows that the suppression of the ST model is not sufficiently strong to describe the data. In addition to this the ST mass function is diverging towards small masses, while we expect a realistic mass function to drop to zero at latest below the free streaming scale.

Fig. 5 compares the measurements of the WDM mass functions from a selection of our highest resolution simulations with the CDM case. We note that, whilst for the case

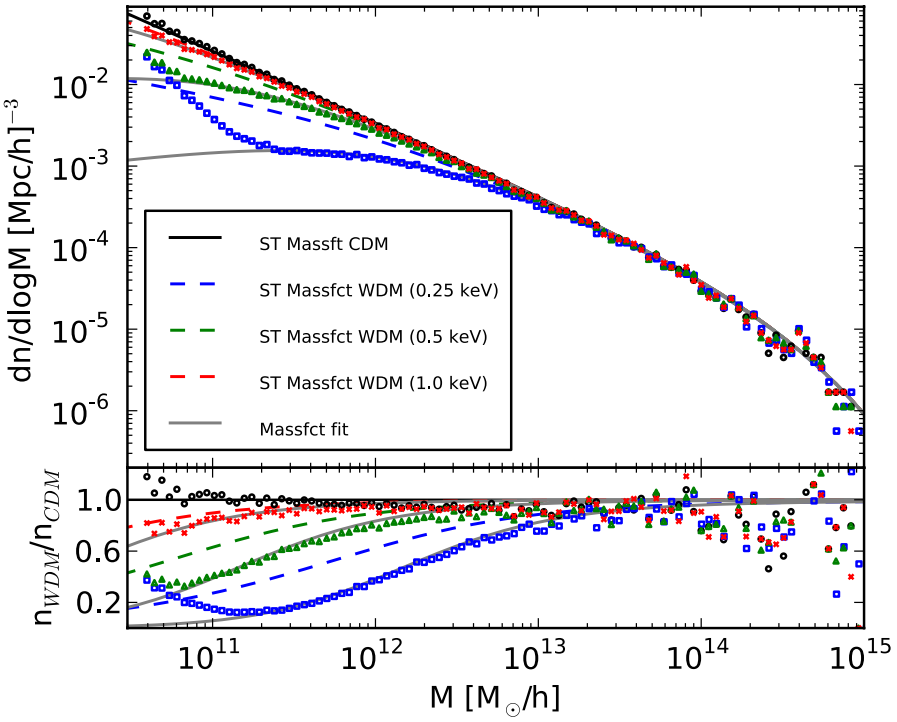

Figure 5. Comparison between the Sheth-Tormen mass functions (black solid line for CDM, colored dashed lines for WDM) and the measurements from the simulations (black circles for CDM, colored squares, triangles and crosses for WDM). The grey solid lines correspond to the mass function fit of Eq. (27).

of CDM the ST model is in very good agreement with the data, the WDM data all lie below the Sheth-Tormen prediction. That is, at least in the mass range above the artificial upturn of the mass function.

Currently, there is no theoretical model that can explain the discrepancy between the CDM and WDM measurements. We shall leave this as an issue for future study. However, it is possible to develop a fitting function that can describe the simulation results to high accuracy. As first noted in Smith & Markovic (2011), if one rescales the mass variable by $M_{\rm fs}$, or equivalently by $M_{\rm hm}$ (cf. §2), then the mass functions for a wide variety of different values of $m_{\rm WDM}$ all appear to fall upon the same locus[4].

In Fig. 6 we show that this scaling also works surprisingly well for the mass function measured from the simulations. We therefore look to fitting the rescaled mass functions. After trying various forms, we found that the function

$$\frac{n^{\rm A}_{\rm WDM}(M)}{n^{\rm ST}_{\rm WDM}(M)} = (1 + M_{\rm hm}/M)^{-\alpha}\,, \qquad (27)$$

which has only one free parameter $\alpha = 0.6$ was able to fit all of our data with an rms error well below five percent.

[4] We find that the locus of theory curves is much tighter than was first noted in Smith & Markovic (2011). This owes to the fact that, they adopted the free-streaming scale of Bardeen et al. (1986); Zentner & Bullock (2003), but used the transfer function of Viel et al. (2005) to generate the actual linear theory power spectra. This slight mismatch led to a slight off-set, which as Fig. 6 shows, is removed when consistent definitions for $M_{\rm fs}$ and $M_{\rm hm}$ are adopted.

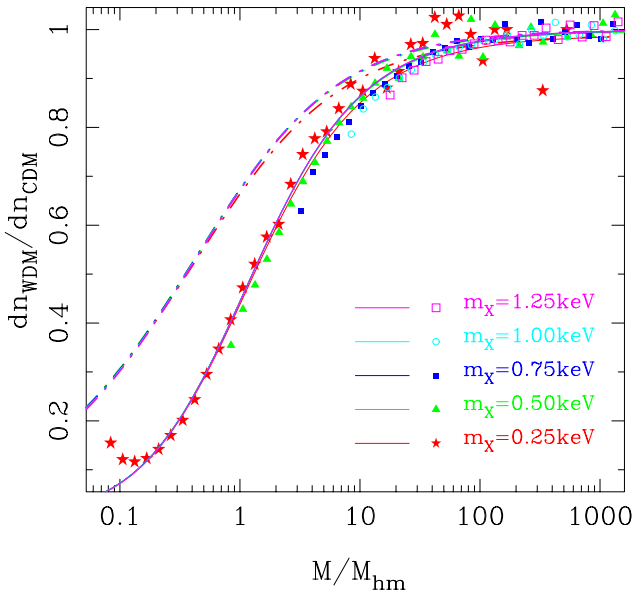

Figure 6. Ratio between the WDM and CDM mass functions, as a function of halo mass, scaled in units of the half-mode mass-scale $M_{\rm hm}$. The measurements from the $N = 1024^3$ suite of WDM simulations are denoted by the point symbols. The dot-dashed lines denote the predictions from the Sheth & Tormen (1999) CDM mass function applied to WDM. The solid lines show the results from the fitting formula of Eq. (27).

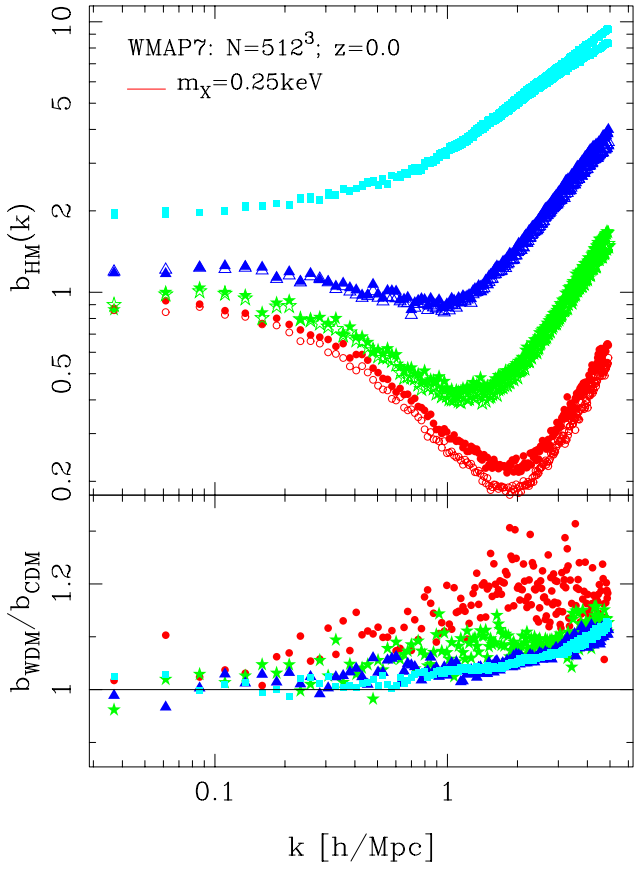

Figure 7. *Top panel*: Comparison of halo bias in the CDM and $m_{\rm WDM} = 0.25$keV WDM model, as a function of wave number. The open and solid points denote the results for CDM and WDM, respectively. Circles, stars, triangles and squares denote results for haloes with masses in the range: $\log_{10}\left(M/[h^{-1}M_\odot]\right) \in \{[12.0, 12.5], [12.5, 13.0], [13.0, 13.5][>13.5]\}$ *Lower panel:* Ratio of the bias in the WDM model with that for the CDM model. For $k > 0.1\,h\,{\rm Mpc}^{-1}$ we see a relative excess signal in the bias of haloes with $M > 10^{12}h^{-1}M_\odot$ in the WDM models. For $k < 0.1\,h\,{\rm Mpc}^{-1}$ the trends are unclear owing to sample variance.

Note that in the above, $n^{\rm ST}_{\rm WDM}$ is the Sheth-Tormen model evaluated for the WDM model in question. The resulting mass functions are plotted as the grey solid lines in Fig. 5. A slightly worse fit may be obtained by using the function,

$$\frac{n^{\rm B}_{\rm WDM}(M)}{n_{\rm CDM}(M)} = (1 + M_{\rm hm}/M)^{-\beta}\,, \tag{28}$$

with $\beta = 1.16$ and this has the advantage that, one only needs to evaluate the CDM mass function and rescale the masses. We note that whilst this paper was being prepared, a similar study was presented by Dunstan et al. (2011), who showed that $n^{\rm B}_{\rm WDM}$ provided a good description of their data but with the slightly higher value $\beta = 1.2$.

Finally, we examined the evolution of the WDM mass-functions up to $z = 1$ and found that Eq. (27) also provides a good description of this data. The simplicity and generality of the fitting function Eq. (27) is surprising and we think that it will be a useful empirical formula.

4.2 Halo bias

We are also interested in understanding how the density fields of dark matter haloes and matter are related in the WDM framework. This relation is usually termed bias, and if we assume that bias is local, deterministic, and linear, then we may write:

$$\delta_{\rm h}(\mathbf{x}|M) = b(M)\delta_{\rm m}(\mathbf{x})\,, \tag{29}$$

where $b(M)$ is the linear bias coefficient, which depends only on the mass of the halo. On using the excursion set formalism and the peak background split argument, one may obtain a prediction for $b(M)$ (Cole & Kaiser 1989; Mo & White 1996; Sheth & Tormen 1999):

$$b^{\rm ST}(\nu) = 1 + \frac{q\nu - 1}{\delta_c(z)} + \frac{2p}{\delta_c(z)\,[1 + (q\nu)^p]}\,, \tag{30}$$

where the parameters p and q are as in Eq. (25). As was shown in Smith & Markovic (2011), if we apply this formula to the case of WDM, then we would expect to see that for $M > M_{\rm hm}$ the bias function is identical to that obtained for CDM. However, for $M < M_{\rm hm}$ we expect to find that the halo bias is increased relative to the CDM case. This occurs due to the fact that ν tends towards a constant value for $M < M_{\rm hm}$ and so $b^{\rm ST}$ becomes constant as well. We again use the simulations to investigate these predictions.

In order to estimate the halo bias, we first sliced the halo distribution into a set of equal number density mass bins. Then, for each mass bin, we estimate the halo and matter auto-power spectra $P^{\rm hh}(k|M)$ and $P^{\rm mm}(k)$, respectively. Our estimator for the bias at each k-mode and in mass bin M_α, can be written:

$$b_i(k_i, M_\alpha) \equiv \sqrt{\frac{P^{\rm hh}(k_i|M_\alpha) - 1/n_{\rm h}(M_\alpha)}{P^{\rm mm}(k_i)}}\,, \tag{31}$$

where $n_{\rm h}(M_\alpha)$ is the number density of haloes for the mass bin α.

Fig. 7 compares the scale-dependence of the halo bias, for several mass bins, and as a function of the wavemode, for the case of CDM (open points) and for the $m_{\rm WDM} = 0.25$ keV WDM model (solid points). Note that here we actually present $b^{\rm hm}(k) \equiv P^{\rm hm}/P^{\rm mm}$, where $P^{\rm hm}$ is defined by the relation $(2\pi)^3\delta_{\rm D}(\mathbf{k}+\mathbf{k}')P^{\rm hm} = \langle\delta_{\rm h}(\mathbf{k})\delta_{\rm m}(\mathbf{k}')\rangle$. In examining the ratio $b_{\rm WDM}/b_{\rm CDM}$, we see that there is increased bias in the WDM case.

We then combine the estimates from each Fourier scale using a standard inverse variance weighted estimator (see e.g. Smith et al. 2007). Also, since, in this case, we are mainly interested in determining the effective linear bias, we only include modes with $k < 0.1\,h\,{\rm Mpc}^{-1}$ (cf. Fig. 7). Fig. 8 presents the linear bias measurements together with the predictions from $b^{\rm ST}(M)$ for a selection of the simulated WDM models. The four panels show the cases: CDM, top left; $m_{\rm WDM} = 1.0$keV, top right; $m_{\rm WDM} = 0.5$keV, bottom left; $m_{\rm WDM} = 0.25$keV, bottom right.

Considering the high mass haloes, we find that the bias estimates for CDM and the WDM models appear to be in reasonable agreement with one another. At lower masses, however, there is a prominent increase in the bias for the WDM models with the lightest particle masses. We have found that, rather than a genuine effect due to WDM initial conditions, this boost appears to be a manifestation of

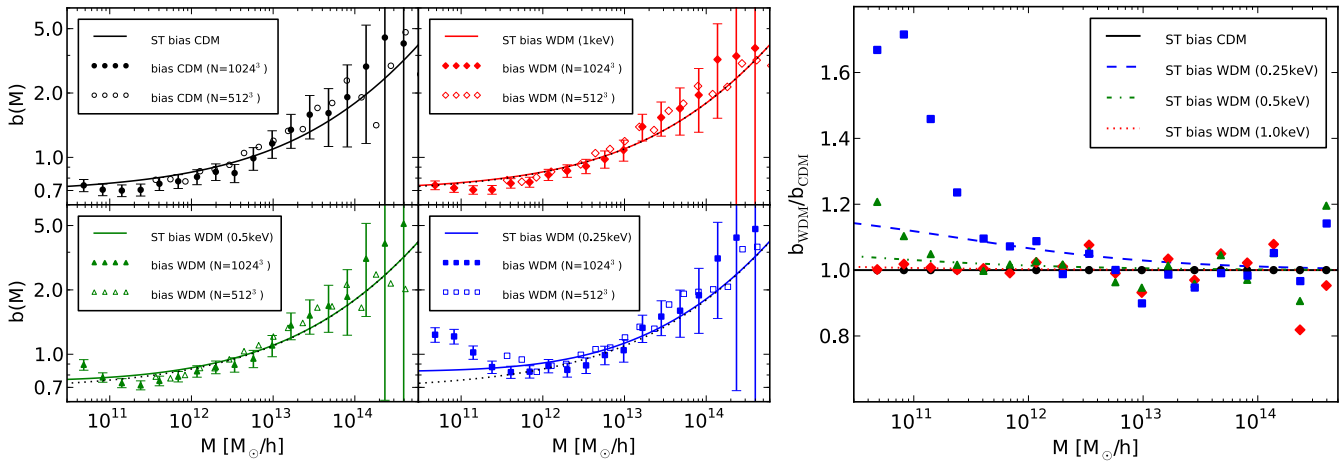

Figure 8. Left panel: Linear halo bias for CDM (top left) and WDM (top right: 1.0 keV, bottom left: 0.5 keV and bottom right: 0.25 keV). The filled and empty dots are measurements from the simulations with $N = 1024^3$ and $N = 512^3$, respectively. Error bars are calculated with an inverse variance wieghted estimator (see Smith et al. 2007). The solid lines correspond to the Sheth-Tormen model prediction of Eq. (30). The linear halo bias of CDM is shown as a black dotted line for comparison. Right panel: Ratios between the WDM and the CDM linear halo bias for the $N = 1024^3$ runs. The error bars have been omitted for clarity.

the artificial halo clumping discussed in §3. This becomes obvious by looking at the bottom-right plot in the panel on the left of Fig. 8, where the upturn of the high resolution simulation (solid blue squares) is shifted with respect to the upturn in the low resolution simulation (empty blue squares). The mass scales of the upturn in the halo bias coincides with the upturn in the mass function (see Fig. 4). Importantly, this means that the halo-halo power spectrum is strongly contaminated by the spurious haloes, even on scales that are considered to be linear. To some extent this is not so surprising, given that below a certain mass scale we are dominated by spurious clumps.

In general, when considering masses above the minimum mass-scale that we trust, rather than an excess bias with respect to $b^{\rm ST}$, we see that the estimates appear to lie slightly below the theoretical prediction at low masses. This has been observed in the CDM framework by Tinker et al. (2010), and it seems to be the case for both our CDM and WDM simulations. However, for the case of $m_{\rm WDM} = 0.25$, we do note that, just above the non-physical upturn, there is a sign that bias in the WDM simulations is larger than in the CDM case. This trend is in qualitative agreement with the Sheth-Tormen prediction for WDM. However, the effect is small and of the order of the error bars and one would need both larger volume and higher resolution simulations to robustly confirm this.

4.3 Anti-bias of the smooth component

We also require the density field of the smooth matter. As for the case of the halo bias, if we assume that this is a linear, deterministic function of the matter density, then we may write the simple expression:

$$\delta_{\rm s}(\mathbf{x}) = b_{\rm s}\delta(\mathbf{x}) \ , \tag{32}$$

where $b_{\rm s}$ is the smooth bias parameter. As shown in Smith & Markovic (2011), this can be calculated using a mass conservation argument, and one finds:

$$b_s = \frac{1}{1-f}\left[1 - \frac{1}{\bar\rho}\int dM M n(M) b_1(M)\right] \leqslant 1 \tag{33}$$

Unlike the halo bias, which is mass dependent, the linear bias of the smooth component stays constant over all scales. In consequence, the smooth component of the power spectrum is directly proportional to the linear matter power spectrum. We shall leave it for future study to establish the veracity of this expression.

4.4 Density profiles

Over the years, extensive numerical work has shown that, for the case of the CDM model, the density profiles of dark matter haloes are reasonably well characterized by the NFW profile (Navarro et al. 1997; Moore et al. 1999; Diemand et al. 2004; Springel et al. 2008; Stadel et al. 2009). This has the universal form:

$$\frac{\rho(r)}{\bar\rho} = \frac{\delta_{\rm s}}{y(1+y)^2} \ ; \qquad y \equiv \frac{r}{r_{\rm s}} \ , \tag{34}$$

where the two parameters $\delta_{\rm s}$ and $r_{\rm s}$ represent a characteristic overdensity and scale radius. The mass of each halo can be determined by simply summing up the number of particles in a given object and multiplying by the particle mass. We can connect this to the virial radius through the relation:

$$M_{\rm vir} = \frac{4\pi}{3}\bar\rho\Delta_{\rm vir}r_{\rm vir}^3, \tag{35}$$

where $r_{\rm vir}$ and $\Delta_{\rm vir}$ are the virial radius and overdensity, respectively. The value of $\Delta_{\rm vir}$ is typically chosen to denote the overdensity for virialization, and here we adopt the value $\Delta_{\rm vir} = 200$ (e.g. see Sheth & Tormen 1999). However, the halo mass $M_{\rm vir}$ can also be obtained by integrating the density profile up to $r_{\rm vir}$, which gives

$$M_{\rm vir} = 4\pi\bar\rho\delta_{\rm s}r_s^3\left[\log(1+c_{\rm vir}) - c_{\rm vir}/(1+c_{\rm vir})\right] \ , \tag{36}$$

where we have introduced the concentration parameter, defined as $c_{\rm vir} \equiv r_{\rm vir}/r_{\rm s}$. On equating Eqs (35) and (36) we find that

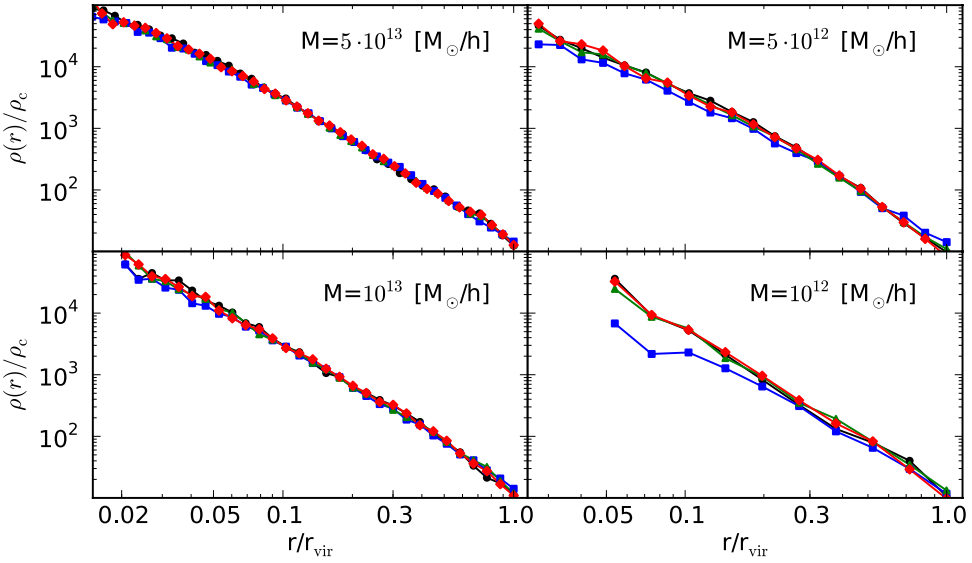

Figure 9. Measurement of the halo profiles for CDM (black) and WDM (blue: 0.25 keV, green: 0.5 keV, red: 1.0 keV) for different halo masses. The profiles of each mass bin are coming from a randomly chosen halo, which is identified in the CDM and all WDM simulations.

$$\delta_{\rm s} = \frac{c^3 \Delta_{\rm vir}/3}{[\log(1+c_{\rm vir}) - c_{\rm vir}/(1+c_{\rm vir})]} \ . \tag{37}$$

This means that the original parameters $\{\delta_{\rm s}, r_{\rm s}\}$ of the NFW profile can be replaced by $\{M_{\rm vir}, c_{\rm vir}\}$. Thus, given a simulated halo of mass $M_{\rm vir}$, the model has one free parameter, the concentration parameter $c(M)$

Fig. 9 shows the density profiles of several randomly chosen haloes of different masses, for the case of CDM (black connected points). We have matched these objects with their counterpart haloes in the our standard set of WDM models and their profiles are also plotted (coloured connected points). While these profiles, on this logarithmic plot, all appear virtually indistinguishable for high masses, there does appear to be a net flattening off in the inner radius for the galaxy mass haloes in the WDM model with $m_{\rm WDM} = 0.25$keV. One important point that can not be easily gleaned from the figure, is that there is an overall reduction in the masses of all the smaller haloes. As we will see, this will have important consequences when we characterize the $c(M)$ relation in the next section.

Earlier work on this topic by Moore et al. (1999) found that there was almost no perceptible difference between CDM haloes and haloes that formed from CDM initial conditions that had no small scale power below a certain scale. Subsequent work by Avila-Reese et al. (2001) and Colín et al. (2008), with a more careful treatment of the WDM transfer function, have shown more significant differences. However, in this case they were exploring models that were closer to HDM than WDM. We therefore conclude that our results are broadly consistent with all of these findings.

One further point, is that for this small sample, we see no visible signs of the formation of a constant density core. This is in agreement with work Villaescusa-Navarro & Dalal (2011). Adding thermal velocities into the simulations could in principle lead to the formation of a constant density core through the Tremaine & Gunn (1979) limit on the fine grained phase space. However, the thermal velocities of WDM cool down with the expansion of space and are already very small during the epoch of structure formation. Thus, if cores are induced, we expect that they will lie below the resolution limit of our simulations (see Kuzio de Naray et al. 2010, for more discussion of this).

In summary, NFW profiles remain a valid approximation for density profiles in our WDM simulations, given our spatial resolution and choice of $m_{\rm WDM}$.

4.5 Concentration-mass relation

As shown in the previous section the NFW model can be fully characterized by specifying the concentration mass relation. We now explore this for the case of WDM.

In the CDM model, $c_{\rm vir}$ has been shown to be a monotonically decreasing function of $M_{\rm vir}$ (Navarro et al. 1997; Bullock et al. 2001; Macciò et al. 2007, 2008; Neto et al. 2007; Prada et al. 2011). One explanation for this, is that owing to the fact that haloes of different mass form at different times, they are therefore exposed to different background densities at collapse and this influences the final core overdensity. A denser background during collapse leads to generally higher concentrations. These ideas were encapsulated into a simple model for halo concentration by Bullock et al. (2001):

$$c_{\rm vir} = K(z_c+1)/(z+1) \ , \tag{38}$$

where z_c is the redshift of collapse. This can be obtained by solving the relation $\sigma(M_*, z) = 1.686$, where $M_* \equiv F M_{\rm vir}$, is defined to be a constant fraction of the virial mass. The two constants K and F must be calibrated using numerical simulations, and for our CDM simulations we found that $K = 3.4$ and $F = 0.001$ provided a good fit to the data. However, we note that the above arguments are only qualitatively correct, since, as first pointed out by Bullock et al. (2001), there is a large scatter between $c_{\rm vir}$ and $M_{\rm vir}$. This can, in part, be traced to the varying accretion histories and large-scale environments of different haloes of the same final mass.

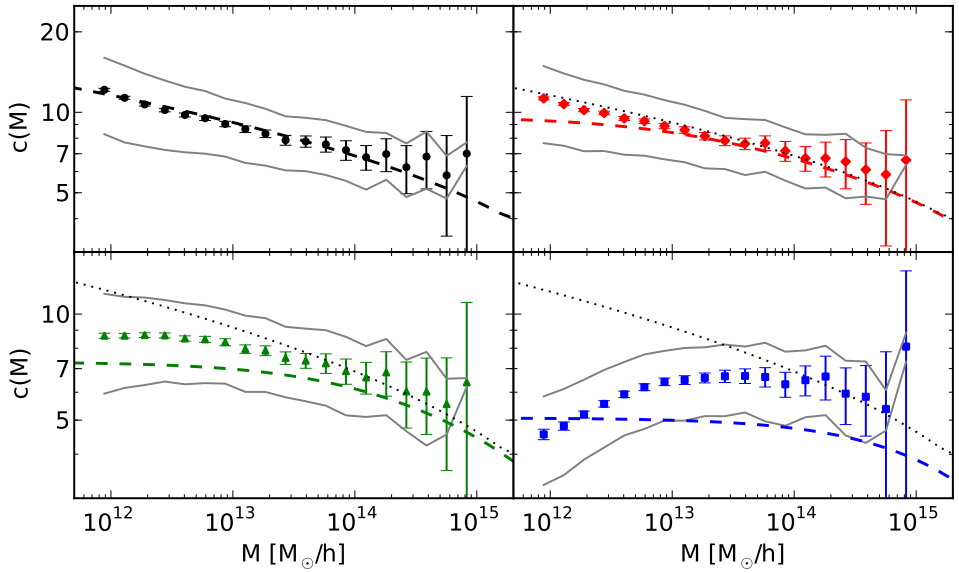

Figure 10. Concentration to mass relation for CDM (top left) and for WDM with $m = 1$ keV (top right), $m = 0.5$ keV (bottom left) and $m = 0.25$ keV (bottom right). The colored symbols denote the median concentrations, while the dashed lines correspond to the Bullock model with $F = 0.001$ and $K = 3.4$. For comparison the Bullock model for CDM has been added to the WDM plots in form of a black dotted line. The gray lines are the 1-σ contours.

Turning to the WDM case, if we directly apply the Bullock model, but using the WDM linear power spectrum, then we find a suppression and a flattening of halo concentrations for masses $M < M_{\rm hm}$. Similar to the mass function, this arises due to the fact that $\sigma(M)$ saturates to a constant value for masses approaching $M_{\rm fs}$. We have tested these predictions, by estimating the concentration parameters for all *relaxed* haloes in our CDM and WDM simulations that contain more than $N = 500$ particles (for full details of the method that we employ see Macciò et al. 2007, 2008).

Fig. 10 shows the measured halo concentrations as a function of mass for a selection of the highest resolution CDM and WDM simulations. The gray solid lines correspond to the 1-σ contours of the measurements, indicating a considerable spread in the concentration-mass relation. The large solid symbols denote the median, with the errors being computed on the mean, i.e. we use $\sigma/\sqrt{N_i}$, where N_i are the number of haloes in the ith mass bin. The dashed lines denote the predictions from the Bullock model. For the CDM case it works reasonably well, especially with our modified parameters $\{K, F\}$. However, the model shows the wrong qualitative behavior for the WDM scenario: whilst the curve for the Bullock model always flattens out towards low masses, we see that for the cases of the lighter WDM particles, there is a turnover in the relation. This turnover in the $c_{\rm vir}$–M relation at low masses is important, as it indicates the end of hierarchical collapse and the emergence of a period of top-down structure formation. As a test of these results we performed additional WDM runs with the `Gadget-2` gravity code (Springel 2008) and we observe the same turnover in this independent set of simulations.

In order to model the $c_{\rm vir}$–M relation for WDM, we shall adapt the Bullock model. As in the case of the mass function, we do this by introducing a correction function described by the relation:

$$\frac{c_{\rm WDM}(M)}{c_{\rm CDM}(M)} = \left(1 + \gamma_1 \frac{M_{\rm hm}}{M}\right)^{-\gamma_2} \quad , \tag{39}$$

where we have again rescaled the halo mass by $M_{\rm hm}$. Least-squares optimization of the free-parameters gives: $\gamma_1 = 15$ and $\gamma_2 = 0.3$.

In Fig. 11 we compare the fitting function (gray solid lines) with the results from the simulations. The parametric relation describes the $c_{\rm WDM}$–M relation with a precision of better than 10% (the fit appears less satisfying for the case $m_{\rm WDM} = 0.5$keV, but only for the lower mass bins). Interestingly, the value $\gamma_1 \sim 10$, informs us that the $c(M)$ relation is sensitive to the presence of WDM for mass scales one order of magnitude larger than for the mass function. As we will see in the next section, this will be important for modeling the nonlinear power spectrum on small scales.

5 NONLINEAR POWER SPECTRUM

5.1 Comparison with existing models

In Fig. 12 we show the nonlinear matter power spectra estimated from our highest resolution CDM and WDM simulations. One can see that for $k \leqslant 1\,h\,{\rm Mpc}^{-1}$, there appears to be no obvious difference between the CDM and WDM models under consideration. This is in stark contrast with the initial linear theory power spectra (cf. Fig. 2), which show considerable damping for the same scales. Clearly nonlinear evolution has regenerated a high-k tail to the power spectrum (cf. White & Croft 2000). At higher wavenumbers $k > 1\,h\,{\rm Mpc}^{-1}$, the situation is more interesting, and we see that the measured WDM power spectra are suppressed with respect to the CDM spectrum. The bottom panel of Fig. 13 quantifies this suppression in greater detail. Here we see that at $k \sim 10\,h\,{\rm Mpc}^{-1}$ there is a 20% suppression in power for the case of $m_{\rm WDM} = 0.25$keV and this drops to $\sim 2\%$ for

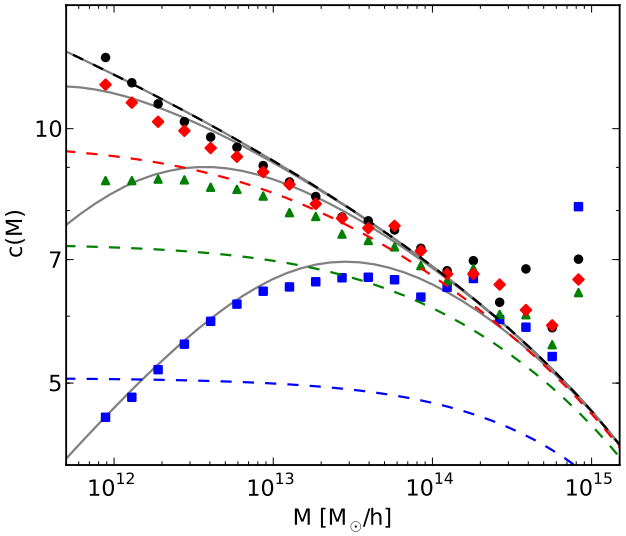

Figure 11. Same as Fig. 10 but measurements are superimposed on one another and without error bars. The additional gray lines illustrate the fitting function from Eq. (39).

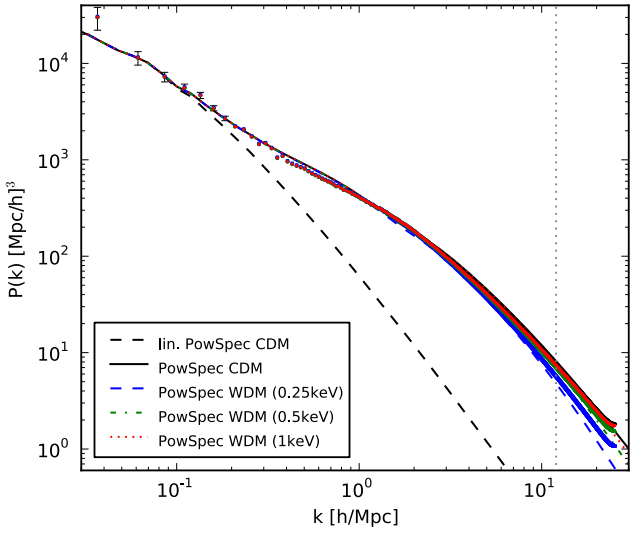

Figure 12. Nonlinear power spectra from the simulations (dots) and from the original halo model (lines), developed by Smith & Markovic (2011). Black corresponds to CDM and color to WDM (red: 1 keV, green: 0.5 keV, blue: 0.25 keV). The vertical gray dots indicate half the Nyquist frequency.

the case $m_{\rm WDM} = 1$keV. The small difference between CDM and WDM at large scales ($k \lesssim 1$) is coming from a shift in the amplitude of the linear power spectrum, fixed with the same σ_8.

We now explore whether the halo model approach, described in §2.2, can accurately reproduce our results for the WDM power spectra. In the original WDM halo model calculation of Smith & Markovic (2011), all of the model ingredients (mass function, density profiles and halo bias relation), were obtained by assuming that the CDM relations also applied to the WDM case, provided one computes them using the appropriate linear power spectrum. The results of this approach are presented in Figs 12 and 13 as the coloured line styles.

In Fig. 13 we see that the halo model of Smith & Markovic (2011) under-predicts the WDM power spectra by roughly $\sim 10\%$. This is reasonably good, considering the assumptions that went into the model. This discrepancy was also noted in the study of Viel et al. (2011). In the bottom panel of Fig. 13 we have also compared our nonlinear power spectra with the predictions from the fitting formula presented in Viel et al. (2011). For scales $k < 10\,h\,{\rm Mpc}^{-1}$ we find that this fitting function provides an excellent description of our data. However, for $k > 10\,h\,{\rm Mpc}^{-1}$ we find discrepancies, especially for the case $m_{\rm WDM} = 0.25$keV. Whether this is a genuine failing of the fitting formula is not clear, since this scale coincides with $\sim k_{\rm Ny}/2$, where $k_{\rm Ny} = \pi N_{\rm grid}/L$ is the Nyquist frequency and we have used $N_{\rm grid} = 2048$.

In summary, we find that the original halo model overestimates the suppression of power due to WDM. This is not too surprising, since we have seen in the previous section that the original approximations for the halo mass function and concentrations turn out to be insufficient descriptions of the simulation data.

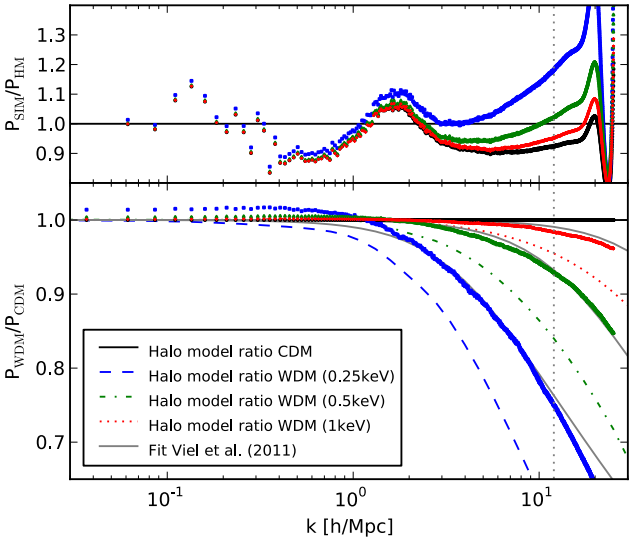

Figure 13. *Top panel*: Ratio of the simulated matter power spectra with respect to the halo model predictions as a function of wavenumber. Different coloured symbols denote the CDM and a selection of the WDM models. *Bottom panel:* Ratio of the WDM and CDM power spectra as a function of wavenumber. Points denote the results from the ratios of simulation data; lines denote the halo model results. The gray solid lines correspond the fitting function from Viel et al. (2011).

5.2 Towards an improved WDM halo model

We now explore whether the halo model predictions can be improved by employing our better fitting functions for $n(M)$ and $\rho(r|M)$. Before making a final prediction for $P(k)$, we first examine how each modification affects the predictions individually.

If we implement our correction for the WDM mass function in the $P(k)$ predictions, then, since the abundance of

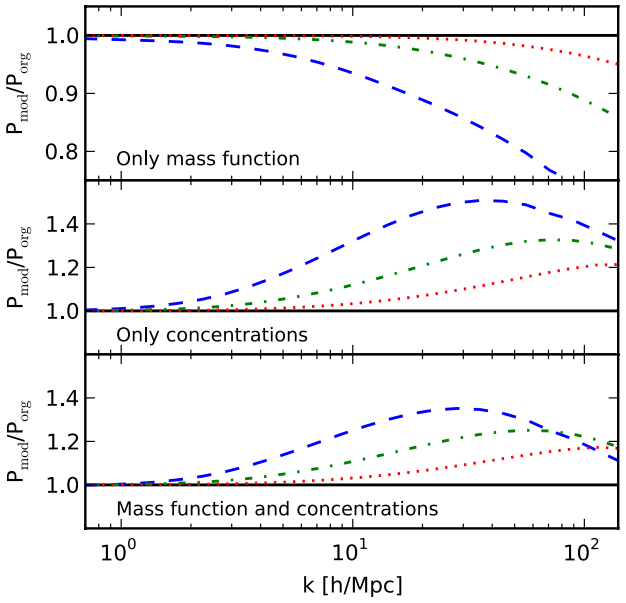

Figure 14. Ratio between modified versions of the halo model and the original version from Smith & Markovic (2011). The black solid line corresponds to CDM and the colored lines to WDM (red dotted: 1 keV, green dashed-dotted: 0.5 keV and blue dashed; 0.25 keV). Top panel: only modification of the mass function. Middle panel: only modification of the concentration-mass relation. Bottom panel: modification of mass function and concentration-mass relation.

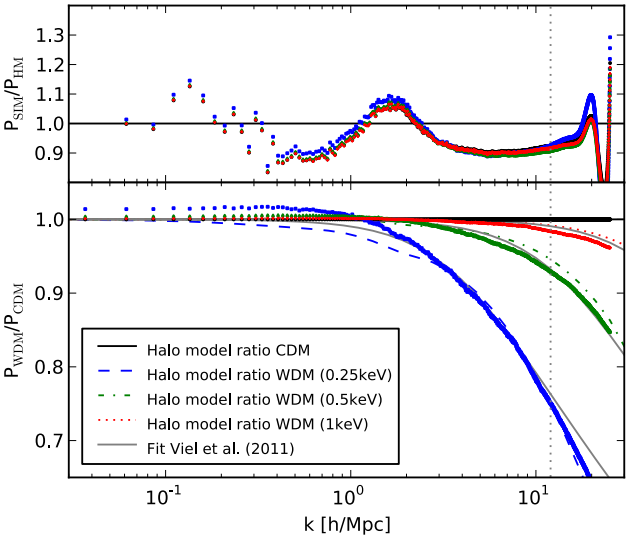

Figure 15. Nonlinear power spectra from the simulations (dots) and from the fully corrected halo model (lines), including the fits for the mass function and the concentrations. The labeling is the same than in Fig. 12. The error of the halo model compared to the simulations has dropped below 10 percent (top pannel), the error on the ratio between WDM and CDM has dropped well below 5 percent (bottom pannel).

small haloes is additionally suppressed with respect to the predictions of $n^{\rm ST}$ for WDM, we should expect that there is an even stronger suppression in $P(k)$. This conjecture is confirmed in the top panel of Fig. 14, which presents the ratio between the halo model with our modified mass function and the original one. We clearly see that the ratio always remains below unity. Somewhat surprisingly, we also note that a $\sim 50\%$ change in the abundance of $10^{12}h^{-1}M_\odot$ haloes, leads to a relatively small change, $\lesssim 10\%$, in the power spectrum at $k \lesssim 10\,h\,{\rm Mpc}^{-1}$.

Next, if we instead implement our improved $c_{\rm vir}(M)$ relation, then we find that this has a more significant impact on the spectra. The central panel of Fig. 14 shows the ratio between the halo model with the modified concentrations and the original one. We find that the suppression of the halo concentrations leads to a $\sim 50\%$ boost in the power for $k \sim 40\,h\,{\rm Mpc}^{-1}$.

The lower panel of Fig. 14 shows the combined behavior of both corrections. The ratio between the fully modified halo model and the original one remains larger than unity. Thus, combination of the modified $n(M)$ and $c(M)$, leads to halo model predictions that have relatively more small-scale power.

Finally, in Fig. 15 we present the comparison between our improved halo model and the nonlinear power spectra from the simulations. The top panel presents the ratios between the simulation data and the halo model predictions. The bottom panel shows the ratios of the WDM and CDM results for both the simulations and our modified halo model. Considering $k \gtrsim 3\,h\,{\rm Mpc}^{-1}$, whilst our modified halo model still has some problems predicting the overall absolute value of $P(k)$, the relative changes between the WDM models and CDM are almost exactly predicted, they being accurate to better than $\sim 2\%$ down to $k \sim 10\,h\,{\rm Mpc}^{-1}$. For scales beyond $k \gtrsim 10\,h\,{\rm Mpc}^{-1}$ we see that the halo model also matches the simulations very well. However, again we note that these scales are beyond $k_{\rm Ny}/2$ and so one might worry about aliasing effects. For $k \lesssim 3$ the error is below about 5%, this scales are however suffering from the difficulties in calculating $P^{\rm c}_{\rm hh}$, descirbed in §2.2.

In summary, we conclude that our modified halo model is able to reproduce nonlinear WDM power spectra with the same accuracy as can currently be achieved for CDM.

6 CONCLUSION

In this paper we have explored nonlinear structure formation in the WDM cosmological model, through a large suite of cosmological N-body simulations and through the halo model. The study was done for a set of fully thermalized WDM models with particle masses in the set $m = \{0.25, 0.5, 0.75, 1.0, 1.25\}$keV. These masses range from purely pedagogical models, towards more realistic scenarios for the dark particle.

For the simulations we chose a box size $L = 256\,h^{-1}$Mpc, which was small enough to resolve both the small scales, where WDM effects play an important role, and the large scales, which are required for correct linear evolution of the box-modes. All models were simulated with $N = \{256^3, 512^3, 1024^3\}$ particles. This was done in order to disentangle physical effects from numerical ones.

In the original halo model calculation for WDM by Smith & Markovic (2011), it was shown that in order to

make robust predictions, one requires good understanding of dark matter halo profiles, the mass function and halo bias. In this work we performed a detailed study of all of these ingredients. Our findings can be summarized as follows:

(i) Mass function: Below a certain scale, the WDM mass function is suppressed with respect to CDM. This suppression is considerably stronger than that obtained by simply applying the Sheth-Tormen approach together with the linear power spectrum of WDM. In agreement with Smith & Markovic (2011), we found that the mass functions for the different WDM models could be transformed into a single locus of points. This was achieved by taking the ratio of the WDM mass function with that for CDM, and then rescaling the masses by $M_{\rm hm}$ (or equivalently $M_{\rm fs}$). We used a fitting function similar to that proposed in Dunstan et al. (2011) to link the Sheth-Tormen mass function to the measured one. The fitting function, which has only one free parameter, was able to reproduce all of the data with an accuracy of a few percent. We also found a strong boost in the mass function at very small mass scales. We showed that this was consistent with artificial halo formation around the initial particle lattice (cf. Wang & White 2007).

(ii) Halo bias: We measured the linear halo bias, using the four largest modes in our simulations. For smaller mass haloes, we found a small enhancement of the bias in WDM simulations, which was qualitatively consistent with the predictions of Smith & Markovic (2011). However, owing to the simulation box being too small, we were unable to quantify this more robustly. At very small masses we found a prominent boost in the bias. We found that this was again a sign of artificial halo formation.

(iii) Density profiles: In the CDM model, the density profiles of dark matter haloes can be characterized by an NFW profile, with a monotonically decreasing concentration-mass relation. In the WDM scenario, we have shown that the NFW profile remains valid for the models and resolution limits of our simulations, and we saw no evidence for a central density core. A simple adaption of the CDM concentration-mass relation, would suggest a strong flattening towards small masses. Whilst, we found such a flattening, the measurements in fact revealed a turnover towards smaller masses. This somewhat surprising result may be interpreted as a sign of top-down structure formation. We modelled the mean relation by adapting a fitting formula similar to that for the mass function. Our fit to the $c(M)$ data was good to an accuracy of $\sim 10\%$. Interestingly, we found that the deviations from CDM in the WDM model, appear in the $c(M)$ relation for halo masses one order of magnitude larger than for the mass function.

After analyzing these ingredients in detail and developing new fitting functions for them, we were able to improve the small-scale performance of the WDM halo model. We found that for $k \gtrsim 3\,h\,{\rm Mpc}^{-1}$, we could predict the absolute amplitude of the power spectrum to better than $\sim 10\%$. However, we were able to predict the ratio of the WDM to CDM spectra, at better than $\lesssim 2\%$. This was competitive with the latest fitting formulae (Viel et al. 2011).

One of the many advantages of the halo model based approach, is that we may more confidently extrapolate our power spectra predictions to smaller scales than can be done from a fitting formula, since the model is built on physical quantities. Furthermore, we may also use the model to study the clustering of galaxies (Zehavi et al. 2005). It is hoped that this may lead to a method for constraining WDM models from galaxy clustering studies. Lastly, one further issue for future study, is to establish a better theoretical understanding of what shapes the mass function and halo concentrations in WDM. In particular, in finding the turnover in the concentration mass relation, have we really seen the reversal of bottom-up structure formation. This promises to be an interesting future challenge.

ACKNOWLEDGMENTS

It is a pleasure to thank Donnino Anderhalden, Jürg Diemand and Darren Reed for useful discussions. We also thank Doug Potter and Joachim Stadel for use of their power spectrum code and technical support concerning `pkdgrav`. We thank Roman Scoccimarro for making public his `2LPT` code. AS, RES and BM acknowledge support from the Swiss National Foundation. RES also acknowledges support from a Marie Curie Reintegration Grant and the Alexander von Humboldt Foundation.

REFERENCES

Avila-Reese, V., Colín, P., Valenzuela, O., D'Onghia, E., & Firmani, C. 2001, ApJ, 559, 516
Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, Nature, 311, 517
Bode, P., Ostriker, J. P., & Turok, N. 2001, ApJ, 556, 93
Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440
Bond, J. R. & Szalay, A. S. 1983, ApJ, 274, 443
Boyanovsky, D. 2010, ArXiv e-prints
Boyarsky, A., Iakubovskyi, D., Ruchayskiy, O., & Savchenko, V. 2008, MNRAS, 387, 1361
Boyarsky, A., Lesgourgues, J., Ruchayskiy, O., & Viel, M. 2009a, Journal of Cosmology and Astro-Particle Physics, 5, 12
—. 2009b, Physical Review Letters, 102, 201304
Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R., & Dekel, A. 2001, MNRAS, 321, 559
Bullock, J. S., Kravtsov, a. A. V., & Colín, P. 2002, ApJL, 564, L1
Cole, S. & Kaiser, N. 1989, MNRAS, 237, 1127
Cole, S., Percival, W. J., Peacock, J. A., & The 2dFGRS Team. 2005, MNRAS, 362, 505
Colín, P., Avila-Reese, V., & Valenzuela, O. 2000, ApJ, 542, 622
Colín, P., Valenzuela, O., & Avila-Reese, V. 2008, ApJ, 673, 203
Colombi, S., Dodelson, S., & Widrow, L. M. 1996, ApJ, 458, 1
Cooray, A. & Sheth, R. 2002, Phys. Rep. , 372, 1
Crocce, M., Pueblas, S., & Scoccimarro, R. 2006, MNRAS, 373, 369

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
de Blok, W. J. G., Walter, F., Brinks, E., Trachternach, C., Oh, S.-H., & Kennicutt, Jr., R. C. 2008, Astron. Journal, 136, 2648
de Vega, H. J. & Sanchez, N. G. 2010, MNRAS, 404, 885
de Vega, H. J. & Sanchez, N. G. 2011, ArXiv e-prints
Diemand, J. & Kuhlen, M. 2008, ApJL, 680, L25
Diemand, J., Moore, B., & Stadel, J. 2004, MNRAS, 353, 624
Dodelson, S. & Widrow, L. M. 1994, Physical Review Letters, 72, 17
Dunstan, R. M., Abazajian, K. N., Polisensky, E., & Ricotti, M. 2011, ArXiv e-prints
Ellis, J., Hagelin, J. S., Nanopoulos, D. V., Olive, K., & Srednicki, M. 1984, Nuclear Physics B, 238, 453
Gentile, G., Famaey, B., Zhao, H., & Salucci, P. 2009, Nature, 461, 627
Gorbunov, D., Khmelnitsky, A., & Rubakov, V. 2008, Journal of High Energy Physics, 12, 55
Jungman, G., Kamionkowski, M., & Griest, K. 1996, Phys. Rep. , 267, 195
Kawasaki, M., Sugiyama, N., & Yanagida, T. 1997, Modern Physics Letters A, 12, 1275
Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82
Kolb, E. W. & Turner, M. S. 1990, The early universe. (Addison Wesley, 1990)
Komatsu, E., Smith, K. M., Dunkley, J., & The WMAP Team. 2011, ApJS, 192, 18
Kuzio de Naray, R., Martinez, G. D., Bullock, J. S., & Kaplinghat, M. 2010, ApJL, 710, L161
Lewis, A., Challinor, A., & Lasenby, A. 2000, Astrophys. J., 538, 473
Lovell, M., Eke, V., Frenk, C., Gao, L., Jenkins, A., Theuns, T., Wang, J., Boyarsky, A., & Ruchayskiy, O. 2011, ArXiv e-prints
Macciò, A. V., Dutton, A. A., & van den Bosch, F. C. 2008, MNRAS, 391, 1940
Macciò, A. V., Dutton, A. A., van den Bosch, F. C., Moore, B., Potter, D., & Stadel, J. 2007, MNRAS, 378, 55
Macciò, A. V. & Fontanot, F. 2010, MNRAS, 404, L16
Markovic, K., Bridle, S., Slosar, A., & Weller, J. 2010, ArXiv e-prints
Miranda, M. & Macciò, A. V. 2007, MNRAS, 382, 1225
Mo, H. J. & White, S. D. M. 1996, MNRAS, 282, 347
Moore, B., Quinn, T., Governato, F., Stadel, J., & Lake, G. 1999, MNRAS, 310, 1147
Moroi, T., Murayama, H., & Yamaguchi, M. 1993, Physics Letters B, 303, 289
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Neto, A. F., Gao, L., Bett, P., Cole, S., Navarro, J. F., Frenk, C. S., White, S. D. M., Springel, V., & Jenkins, A. 2007, MNRAS, 381, 1450
Peebles, P. J. E. 1982, ApJL, 263, L1
—. 2001, ApJ, 557, 495
Peebles, P. J. E. & Nusser, A. 2010, Nature, 465, 565
Polisensky, E. & Ricotti, M. 2010, ArXiv e-prints
Prada, F., Klypin, A. A., Cuesta, A. J., Betancort-Rijo, J. E., & Primack, J. 2011, ArXiv e-prints
Press, W. H. & Schechter, P. 1974, ApJ, 187, 425
Salucci, P., Lapi, A., Tonini, C., Gentile, G., Yegorova, I., & Klein, U. 2007, MNRAS, 378, 41
Scoccimarro, R. 1998, MNRAS, 299, 1097
Seljak, U., Makarov, A., McDonald, P., & Trac, H. 2006, Physical Review Letters, 97, 191303
Shaposhnikov, M. & Tkachev, I. 2006, Physics Letters B, 639, 414
Sheth, R. K. & Tormen, G. 1999, MNRAS, 308, 119
Smith, R. E., Desjacques, V., & Marian, L. 2011, PRD, 83, 043526
Smith, R. E. & Markovic, K. 2011, ArXiv e-prints
Smith, R. E., Scoccimarro, R., & Sheth, R. K. 2007, PRD, 75, 063512
Smith, R. E., Peacock, J. A., Jenkins, A., White, S. D. M., Frenk, C. S., Pearce, F. R., Thomas, P. A., Efstathiou, G. & Couchman, H. M. P. 2007, MNRAS, 341, 1311
Springel, V. 2005, MNRAS, 364, 1105
Springel, V., Wang, J., Vogelsberger, M., Ludlow, A., Jenkins, A., Helmi, A., Navarro, J. F., Frenk, C. S., & White, S. D. M. 2008, MNRAS, 391, 1685
Stadel, J., Potter, D., Moore, B., Diemand, J., Madau, P., Zemp, M., Kuhlen, M., & Quilis, V. 2009, MNRAS, 398, L21
Stadel, J. G. 2001, PhD thesis, AA(UNIVERSITY OF WASHINGTON)
Swaters, R. A., Madore, B. F., van den Bosch, F. C., & Balcells, M. 2003, ApJ, 583, 732
Tegmark, M., , & The SDSS Team. 2006, PRD, 74, 123507
Tikhonov, A. V., Gottlöber, S., Yepes, G., & Hoffman, Y. 2009, MNRAS, 399, 1611
Tinker, J. L., Robertson, B. E., Kravtsov, A. V., Klypin, A., Warren, M. S., Yepes, G., & Gottlöber, S. 2010, ApJ, 724, 878
Tremaine, S. & Gunn, J. E. 1979, Physical Review Letters, 42, 407
Viel, M., Lesgourgues, J., Haehnelt, M. G., Matarrese, S., & Riotto, A. 2005, PRD, 71, 063534
Viel, M., Markovic, K., Baldi, M., & Weller, J. 2011, ArXiv e-prints
Villaescusa-Navarro, F. & Dalal, N. 2011, Journal of Cosmology and Astro-Particle Physics, 3, 24
Wang, J. & White, S. D. M. 2007, MNRAS, 380, 93
White, M. & Croft, R. A. C. 2000, ApJ, 539, 497
Zavala, J., Jing, Y. P., Faltenbacher, A., Yepes, G., Hoffman, Y., Gottlöber, S., & Catinella, B. 2009, ApJ, 700, 1779
Zehavi, I., Zheng, Z., Weinberg, D. H., Frieman, J. A., Berlind, A. A., Blanton, M. R., Scoccimarro, R., Sheth, R. K., Strauss, M. A., Kayo, I., Suto, Y., Fukugita, M., Nakamura, O., Bahcall, N. A., Brinkmann, J., Gunn, J. E., Hennessy, G. S., Ivezić, ***v***Z., Knapp, G. R., Loveday, J., Meiksin, A., Schlegel, D. J., Schneider, D. P., Szapudi, I., Tegmark, M., Vogeley, M. S., York, D. G., & SDSS Collaboration. 2005, ApJ, 630, 1
Zentner, A. R. & Bullock, J. S. 2003, ApJ, 598, 49

5

Paper III: Stability of tidal streams

In the last part of the thesis we are exploring the gravitational collapse within a cylindrical dark matter distribution. The evolution of linear perturbations in an infinitely long cylinder is qualitatively different from the case of a homogeneous distribution, and peculiar characteristics of structure formation such as the Jeans criterion, the linear growing modes, and nonlinear collapse have to be worked out independently.

The following manuscript is focused on the gravitational stability of tidal streams, emerging from dark matter substructures on an orbit around a galaxy. The model of a cylindrical distribution is thereby used to show that tidal streams do not self gravitate because their internal velocity dispersion is too high and because they are expanding with time. The modes of perturbation are either oscillating or they are freezing out after an unsubstantial period of growth.

The paper has been published 2011 in the Monthly Notices of the Royal Astronomical Society [42].

Mon. Not. R. Astron. Soc. **000**, 1–9 (2011) Printed 24 June 2011 (MN LATEX style file v2.2)

On the Stability of Tidal Streams

Aurel Schneider and Ben Moore
Institute for Theoretical Physics, University of Zurich, Zurich, Switzerland;
aurel@physik.uzh.ch; moore@physik.uzh.ch

24 June 2011

ABSTRACT
We explore the stability of tidal streams to perturbations, motivated by recent claims that the clumpy structure of the stellar streams surrounding the globular cluster Palomar 5 are the result of gravitational instability. We calculate the Jeans length of tidal streams by treating them as a thin expanding cylinder of collisionless matter. We also find a general relation between the density and the velocity dispersion inside a stream, which is used to determine the longitudinal Jeans criterion. Our analytic results are checked by following the time evolution of the phase space density within streams using numerical simulations. We conclude that tidal streams within our galactic halo are stable on all length scales and over all timescales.

Key words: galaxies: kinematics and dynamics - galaxies: star clusters - methods: analytical

1 INTRODUCTION

Tidal streams are a widespread phenomenon in astrophysics, emerging from star clusters (Grillmair, Freeman, Irwin & Quinn 1995), dark matter subhalos (Diemand, Kuhlen, Madau, Zemp, Moore, Potter & Stadel 2008) or satellites of galaxies (Ibata, Gilmore & Irwin 1994). It is present on all scales from galaxy clusters (Calcneo-Roldan, Moore, Bland-Hawthorn, Malin & Sadler 2000) down to the the very smallest dark matter substructures (Schneider, Krauss & Moore 2010). The gravitationally unbound material forms spectacular long streams that trace the past and future orbit of the host system.

A wide class of tidal streams can be treated as collisionless systems, since they are dominated by stars or dark matter particles - the local relaxation time within the stream is much longer than the age of the Universe. Some streams contain gaseous material and are much more complicated to understand. For example, the oldest example of a 'stream' is the spectacular Magellanic HI stream, trailing well over 100 degrees behind the Magellanic Clouds. Initially modelled as a tidal mass loss feature from the Large Magellanic Clouds (Lin & Lynden-Bell (1977)), an alternative explanation is that it resulted from a more complex gravitational interaction between the Large and the Small Magellanic Cloud prior to their infall in the Milky Way potential (Besla, Kallivayalil, Hernquist, van der Marel, Cox & Kere 2010). However, it may also be the case that this feature is purely hydrodynamical in origin since it contains no stars (Moore & Davis 1994).

Galaxy mergers often create spectacular tidal tails that are somewhat different from the streams we consider in this paper. These streams are rapidly and violently created and they can contain dwarf galaxies aligned along the tails. Barnes & Hernquist (1992) carried out simulations of galaxy mergers and found collapsed objects populating the stellar tails. Therefore they proposed collisionless collapse as the creation mechanism of tidal dwarf galaxies. However, Wetzstein, Naab & Burkert (2007) identified these collapsed objects as numerical artefacts due to insufficient resolution. They rather found that it's the gaseous part of the streams that triggers the collapse which leads to tidal dwarf galaxies.

The dynamics of tidal streams from star clusters and dwarf galaxies in our own halo have been extensively studied to constrain the mass and shape of the Galactic potential (Johnston, Zhao, Spergel & Hernquist 1999; Law, Majewski & Johnston 2009), alternative gravity models (Read & Moore 2005) as well as the orbital history of the satellites (Kallivayalil, van der Marel & Alcock 2006; Lux, Read & Lake 2010). However, the detailed evolution of the internal phase space structure of streams has received less attention (Helmi & White 1999; Eyre & Binney 2010).

Simulations, as well as observational data, show variations in the width and the internal structure of tidal streams. Likewise the density along a stream can vary considerably, the most prominent example being the symmetric streams originating from the globular cluster Palomar 5 with its equally spaced density clumps (Odenkirchen et al. 2001; Odenkirchen, Grebel, Dehnen, Rix & Cudworth 2002). There are different explanations for these clumps such as disc shocking (Dehnen, Odenkirchen, Grebel, & Rix 2004), effects due to the dark matter substructures (Mayer, Moore,

Quinn & Governato 2002; Yoon, Johnston & Hogg 2010) or epicyclic motions in the stellar orbits (Küpper, MacLeod & Heggie 2008; Just, Berczik, Petrov & Ernst 2009).

Another interpretation was given recently by Quillen and Comparetta (Quillen & Comparetta 2010; Comparetta & Quillen 2010), who argued that clumps in streams are the result of longitudinal Jeans instabilities. In their model they describe a tidal stream as an extended static cylinder of stars and they use the results of Fridman & Polyachenko (1984), that infinitely extended cylinders are gravitationally unstable. With an estimated relation for the velocity dispersion and the linear density in the stream, Quillen & Comparetta find a longitudinal Jeans length of several times the stream width. Comparing their results to the observations of Palomar 5, they find agreement between the distance between clumps in the streams and their fastest growing mode of the gravitational instability.

However, their model of a static cylinder does not take into account the expansion that happens due to the diffluence of the stars in the stream. Once in the stream the stars are no longer bound to the cluster, their intrinsic dispersion causes the stream to grow along the orbital direction. Escaping stars also have an intrinsic dispersion, related to the dispersion in the outer cluster region. Another way of understanding the expansion is by considering the velocity difference between the substructure and the outflowing stars, which depends on the tidal radius. Since the tidal radius is shrinking with time, stars that leave the cluster at later times are slower than stars that left before and this leads to the expansion of the stream. In reality the situation is even more complicated. The stream length is actually oscillating during one orbit, being stretched at pericenter and compressed at apocenter. The linear expansion only acts on average over several orbital periods. Therefore for short timescales, the periodic oscillating effect must be taken into account.

This paper is structured as follows: In section 2 we construct a simplified model for a tidal stream and we find a relation between the stream density and its velocity dispersion. Sections 3 and 4 are dedicated to the study of the stream stability, where we first derive the linearised equation of perturbations and then look at the one dimensional collapse along the stream direction. In section 5 we take a critical look at our model by comparing with the detailed dynamics of streams using N-body simulations. The orbital oscillation of the stream length and its influence on collapse are considered. Finally we give our conclusions in section 6.

2 MODELLING A TIDAL STREAM

The general case of a streaming cluster is a problem of many particle dynamics that can be solved self-consistently with simulations. Analytical statements can be made by considering a model with simplifying assumptions. Thereby one has to be careful to avoid over-simplification. We model a tidal stream as a self gravitating cylinder of collisionless matter with an expansion in the direction of the cylinder axis. For the cluster as well as for the host we choose isothermal spheres so that we can use the simplifying relations

$$\frac{r_t}{R} \sim c\left(\frac{m}{M}\right)^{1/3} \sim c^{3/2}\left(\frac{\sigma_{cl}}{\sigma_{gal}}\right), \qquad GM = 2\sigma_{gal}^2 R, \tag{1}$$

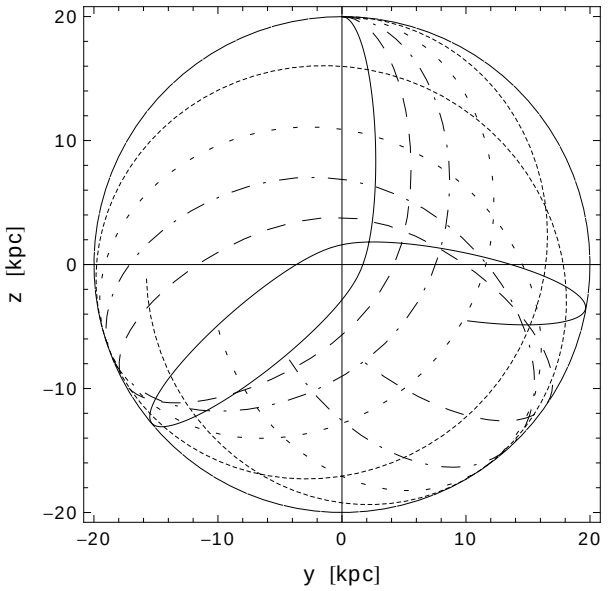

Figure 1. Orbits of the star clusters in our simulations. The eccentricity is given in terms of the parameter b defined as $\dot{R} = bV$ (with $b = 0$ for a circular orbit and $b = 1$ for a radial infall). In increasing eccentricity: continuous ($b = 0$), narrow-dotted ($b = 0.14$) broad-dotted ($b = 0.34$), dashed-dotted ($b = 0.54$) dashed ($b = 0.74$) and continuous ($b = 0.88$).

where m, M and σ_{cl}, σ_{gal} are the masses respectively the velocity dispersions of the cluster and the host. The distances r_t and R are the tidal radius of the cluster and the orbital radius to the host. For an isothermal sphere the correction factor $c \sim 0.8$ (Binney & Tremaine 2008).

Whilst many systems can be reasonably well described by an isothermal potential over the scales of interest, our results would not apply to systems orbiting within very different potentials. For example, a star cluster within a constant density potential would not even produce streams. However, for Palomar 5 and for many of the streams in our Galactic halo, or within galaxy clusters, an isothermal potential is a good approximation over the range $0.01-0.5R_{virial}$ (Klypin, Zhao & Somerville 2002).

In order to test the basic assumptions of our model we perform simulations of a star cluster orbiting with different eccentricities within an isothermal host potential. The velocity dispersion is chosen to be $\sigma_{cl} = 4$ km/s for the cluster and $\sigma_{gal} = 200$ km/s for the host. Every simulation starts with the star cluster at a radius of 20 kpc and we choose different perpendicular initial velocities from 283 km/s for a circular orbit to 50 km/s for the most eccentric orbit. The global potential is a fixed analytic potential whilst the star cluster is modelled using 2×10^5 stars, set up in an equilibrium configuration at the starting position (Zemp, Moore, Stadel, Carollo & Madau 2008). The evolution is followed using the N-body code PKDGRAV (Stadel 2001), adopting high precision parameters for the force accuracy. The softening length of the star particles is $\epsilon = 0.005$ kpc. All simulated orbits are illustrated in Fig. 1.

There are two mechanism responsible for the stream growth, on the one hand, the outflow of matter leaving the cluster with a certain velocity difference ΔV and on the other hand the stream expansion due to diffluence of

the initial dispersion. The expansion velocity w is given by $w \sim 2\sigma_{cl}$, which corresponds to the diffluence velocity of a bunch of particles leaving the cluster at the same time. Using conservation of angular momentum L leads to the velocity difference ΔV:

$$L = RV\sin\theta = (R + r_t)(V - \Delta V)\sin\theta \quad \Rightarrow \quad \frac{\Delta V}{V} \sim \frac{r_t}{R} \quad (2)$$

Here we have assumed $R \gg r_t$ and $V \gg \Delta V$. In an isothermal potential the value of V must be somewhere between $V_R = (4/\pi)^{1/2}\sigma_{gal}$ and $V_c = 2^{1/2}\sigma_{gal}$, the radial and circular velocities. Using (1) we therefore obtain $\Delta V \sim \sigma_{cl}$ as well as

$$w \sim 2\Delta V. \quad (3)$$

Physically this means that all particles belonging to the stream at t_0 will be distributed over the entire stream length at all time $t > t_0$. Or in other words, even if there is no more outflow from the cluster, the stream always stays attached to the cluster.

The amount of diffluence can be estimated in the simulation by marking particles at a certain time t_0 and looking at where they are found in the stream at $t \gg t_0$. The first image of Fig. 2 shows a cluster at apocenter after one orbit (195 Myr) with the particles of one stream marked in red. In the second image we see the cluster at apocenter after nine orbits (1756 Myr) along with the distribution of the particles marked before. The particles marked at the early time are located throughout the stream at later times, confirming our above statement.

The width of the stream depends on the velocity dispersion σ. A particle with an energy excess during the outflow will be on an orbit with a slightly different eccentricity and will therefore complete a full oscillation within the stream during one orbital time T. The radius corresponding to half of the stream width is then approximately given by

$$r_\perp \sim \frac{1}{2}\sigma T. \quad (4)$$

On the other hand the length of the stream after one orbit is simply

$$l_0 \sim wT \sim 2\sigma T, \quad (5)$$

assuming the approximate relation $\sigma_{cl} \sim \sigma$. After one orbit a single stream should therefore be about twice as long as it is wide. This is the case in all our simulations and can be checked in the first image of Fig 2.

The linear density of a stream is given by the relation

$$\mu = \frac{dm}{dz} = \frac{\dot{m}}{\dot{z}} \sim \frac{\dot{m}}{2\Delta V}, \quad (6)$$

Here we have used $\dot{z} \sim w \sim 2\Delta V$, what results in an additional factor of two compared to a static stream because of the stretching effect of the expansion. The rate of outstreaming matter is estimated to be

$$\dot{m} = \frac{(m_a - m_p)}{T} \sim \frac{2c^{\frac{3}{2}}}{T}\frac{\sigma_{cl}^3}{G\sigma_{gal}}(R_a - R_p) \sim \frac{c^{\frac{3}{2}}}{G}\frac{\sigma_{cl}^3}{\sigma_{gal}}\dot{R}, \quad (7)$$

where we have used the relations (1). The outflow of the matter is averaged over one orbital period. With the relation $\dot{R} = bV$, where the parameter b depends on the cluster orbit (with $b = 0$ for a circular orbit and $b = 1$ for a radial infall), the linear density becomes

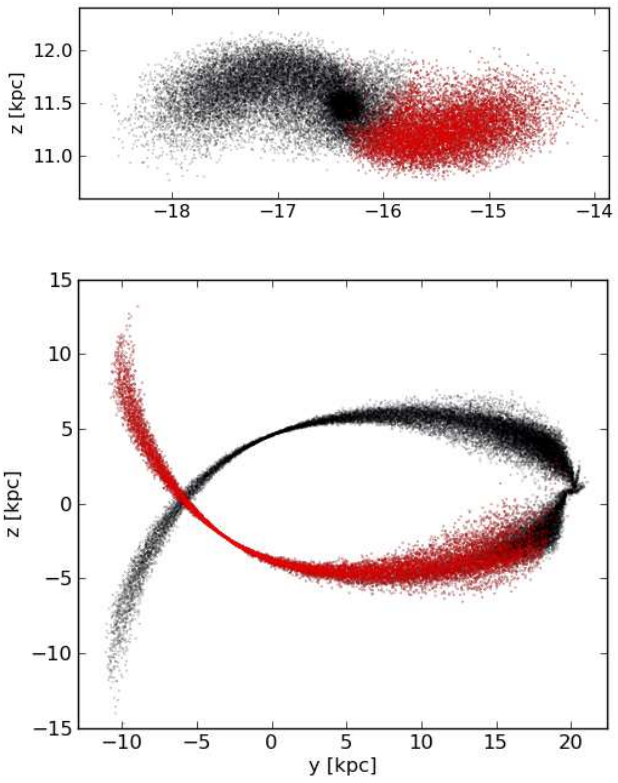

Figure 2. *Simulation of an isothermal cluster with an eccentricity of $b = 0.74$. The image on the top shows the cluster after one orbit where the particles of one stream are marked in red. The image on the bottom shows the cluster after nine orbits with the distribution of the particles marked before.*

$$\mu \sim \frac{c^{3/2}}{2G}\frac{\sigma_{cl}^3}{\sigma_{gal}}\left(\frac{V}{\Delta V}\right)b \sim \frac{\sigma_{cl}^2 b}{2G}, \quad (8)$$

and the Toomre parameter is then given by

$$q \equiv \frac{\sigma^2}{2G\mu} \sim \frac{1}{b}. \quad (9)$$

The smallest value for the Toomre parameter is therefore $q \sim 1$ which corresponds to a radial orbit. The Toomre parameter of relation (9) is four times larger than the one obtained by Quillen & Comparetta (2010), the reason being a factor of two which comes in at equation (6) as well as the averaging of the mass outflow in equation (7). Both effects are directly related to the expansion of the stream, not considered by Quillen and Comparetta.

An independent way to calculate the Toomre parameter is by using the virial theorem for an isothermal sphere, truncated at the tidal radius r_t:

$$\sigma_{cl}^2 = \frac{|W|}{m_{cl}} = \frac{4\pi G}{m_{cl}}\int_0^{r_t} drr\rho(r)M(r) = \frac{4\sigma_{cl}^4 r_t}{Gm_{cl}}, \quad (10)$$

$$\sigma_{cl}^2 = \frac{Gm_{cl}}{4r_t}. \quad (11)$$

Using the approximation $\sigma_{cl} \sim \sigma$ then leads to

$$q = \frac{\sigma^2 l_0}{2Gm_{st}} \sim \frac{1}{8}\frac{m_{cl}}{m_{st}}\frac{l_0}{r_t}. \quad (12)$$

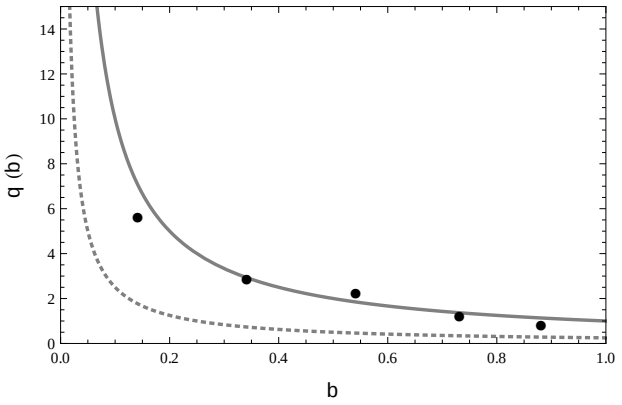

Figure 3. The Toomre parameter as a function of the eccentricity parameter b. The black dots are the measurements from the different simulations. The solid gray line corresponds to equation (9), while the dotted line is the prediction from Quillen & Comparetta (2010).

For the extreme case of a radial orbit $m_{cl} \sim 2m_{st}$ and we obtain $q \sim 1$. This means that for all orbits q must be larger than one, a result that confirms the relation (9) above.

The Toomre parameter can also be determined in the simulations by measuring the velocity dispersion and the linear density. However, it turns out that the dispersion is very difficult to quantify accurately because over one orbital period it strongly fluctuates at any Lagrangian point (for example, around any star). This is due to the oscillation of stream-length and stream width, which happens because the particles in the stream are on nearly free orbits around the host. The velocity dispersion therefore is affected by the number of stars used in its measurement since that changes the region of the stream over which the dispersion is calculated.

In Fig. 3 we plot the Toomre parameter, where the dispersion is measured in the middle of the stream, at apocenter after one orbital period and assuming an isotropic distribution (looking at the ration of tangential to radial velocity dispersions after one orbit we can see that this assumption is approximately valid). We notice that the simulations roughly follow the theoretical prediction which is given by the solid gray line.

Already at that stage of our analysis it becomes clear that the expansion has a strong stabilising effect because it leads to a significant boost of the Toomre parameter. In the next section we will see that the stability of a stream is additionally enforced, since the expanding environment leads to a damping in the the evolution of perturbations.

3 PERTURBATIONS IN AN EXPANDING CYLINDER

In order to find a criterion for the stability, we are now modeling a tidal stream as a non-rotating elongated cylinder of collisionless matter that is linearly expanding in the direction of its long axis. For the expansion we introduce the comoving coordinate $s = az$ with $a(t) = \alpha t$ and set $z = l_0$, where l_0 is the stream length after one orbital period T. The expansion factor then becomes

$$\alpha = \frac{1}{T}. \tag{13}$$

The orbital period is a natural time measure since the outstreaming from the cluster into the tails is mainly happening during the cluster orbit from apo- to pericenter when the tidal radius is shrinking. During the other half of the orbit the tidal radius is growing again and there is nearly no streaming mass loss.

An analytical treatment of the stability of an expanding cylinder is possible either on scales much smaller or much larger than the cylindrical radius. In the former case we can treat the fluid as homogeneous and we therefore get the usual Jeans length

$$\lambda_J^h = \sqrt{\frac{\pi\sigma^2}{G\rho}}. \tag{14}$$

With the relation (9) as well as the linear density $\mu = \pi r_\perp^2 \rho$ we then obtain

$$\frac{\lambda_J^h}{r_\perp} = \sqrt{2\pi^2 q} \sim \sqrt{\frac{2\pi^2}{b}}. \tag{15}$$

Since the eccentricity parameter b is always larger than one, the Jeans length exceeds the radius of the cylinder and we can exclude collapse on scales smaller than $r_\perp$.

However there is still the possibility of collapse in the longitudinal direction of the cylinder on scales larger than $r_\perp$. This is the second analytically treatable case which leads to a very different stability criterion. In order to determine the behaviour of longitudinal perturbations we are now going to derive the equations for the evolution of density perturbations. This is usually done by integrating and linearising the collisionless Boltzmann equation (Peebles 1980). Since we are looking at a thin cylinder, we can assume a phase-space density of the form

$$f(z,p,t) = \begin{cases} a[\rho_b + \rho_1(z,t)]f(p), & r < r_\perp \\ 0, & r > r_\perp \end{cases} \tag{16}$$

Here we have introduced a homogeneous background density ρ_b as well as a first order perturbation ρ_1. An integration of the phase-space density immediately leads to the stream density

$$\rho = \frac{1}{a}\int dp f(z,p,t) = \frac{(\mu_b + \mu_1)}{\pi r_\perp^2} = \frac{1}{\pi r_\perp^2}\frac{\mu_0}{a}(1+D), \tag{17}$$

where $D = \mu_1/\mu_b$ is the dimensionless overdensity.

The evolution of the phase-space density is described by the one dimensional collisionless Boltzmann equation with expanding coordinate

$$\frac{\partial}{\partial t}f(z,p,t) + \frac{p}{a^2}\frac{\partial}{\partial z}f(z,p,t) - \frac{\partial\Phi}{\partial z}\frac{\partial}{\partial p}f(z,p,t) = 0. \tag{18}$$

It is now straightforward to derive the continuity and the momentum equation of the stars in the cylinder. They are given by

$$\partial_t(1+D) + \frac{1}{a}\partial_z\left[\langle v\rangle(1+D)\right] = 0, \tag{19}$$

$$\partial_t\left[a\langle v\rangle(1+D)\right] + \partial_z\Phi(1+D) + \partial_z\left[\langle v^2\rangle(1+D)\right] = 0, \tag{20}$$

where

$$\langle v\rangle = \frac{\int p f dp}{a\int f dp}, \qquad \langle v^2\rangle = \frac{\int p^2 f dp}{a^2\int f dp}. \tag{21}$$

By substituting the derivative of the second equation into the first we finally find the equation of perturbation:

$$\partial_t^2 D + 2\frac{\dot{a}}{a}\partial_t D = \frac{1}{a^2}\partial_z\left[(1+D)\partial_z\Phi\right] + \frac{1}{a^2}\partial_z^2\left[(1+D)\langle v^2\rangle\right]. \quad (22)$$

In order to solve this differential equation we still need to know the potential of a cylinder. The simplest assumption is to take

$$\Phi(r,z,t) = \Phi^{(0)}(r) + \Phi^{(1)}(r,z,t), \quad (23)$$
$$\Phi^{(1)}(r,z,t) = \phi^{(1)}(r,t)e^{ik_0 z}$$

(Fridman & Polyachenko 1984), where k_0 is the comoving wave number in z-direction. The Poisson equation for the zero-order term is simply

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{d\Phi^{(0)}}{dr}\right) = 4\pi G\begin{cases}\rho_0, & r < r_\perp\\ 0, & r > r_\perp\end{cases} \quad (24)$$

with the solution $\Phi^{(0)}(r) = \pi G\rho_0 r^2 + const.$ In contrast to a self gravitating cylinder, a stream is embedded in the dominating potential of the host and the zero order term looks different. However, a dependence of the potential in the z-direction only comes in as a first order effect due to the internal structure of the stream. Therefore we obtain the following Poisson equation at first order

$$\partial_r^2\Phi^{(1)} + \frac{1}{r}\partial_r\Phi^{(1)} - \frac{k_0^2}{a^2}\Phi^{(1)} = \begin{cases}4\pi G\rho_1, & r < r_\perp\\ 0, & r > r_\perp\end{cases} \quad (25)$$

where ρ_1 may vary along the axis of the cylinder. Two independent solutions of this homogeneous differential equation are the modified Bessel equations of first and second kind $I_0[x]$ and $K_0[x]$. The general inner and outer solution are given by

$$\Phi_<^{(1)}(r) = AI_0\left[\frac{k_0 r}{a}\right] + BK_0\left[\frac{k_0 r}{a}\right] - \frac{4\pi G\rho_1 a^2}{k_0^2}, \quad (26)$$

$$\Phi_>^{(1)}(r) = A'I_0\left[\frac{k_0 r}{a}\right] + B'K_0\left[\frac{k_0 r}{a}\right], \quad (27)$$

where the boundary conditions require $B = A' = 0$. With the matching conditions $\Phi_<^{(1)}(r_\perp) = \Phi_>^{(1)}(r_\perp)$ and $\partial_r\Phi_<^{(1)}(r_\perp) = \partial_r\Phi_>^{(1)}(r_\perp)$ we find

$$A = \frac{4\pi G\rho_1 a^2 K_0'\left[k_0 r_\perp/a\right]}{k_0^2 W\left[k_0 r_\perp/a\right]}, \quad W = -\frac{k_0}{a}(I_0K_1 + I_1K_0). \quad (28)$$

We now look at the case of large perturbations in a thin stream ($k_0 r_\perp/a << 1$). In the asymptotic limit we get

$$A \simeq \frac{2G\mu_0}{a}\left[\frac{2a^2}{(k_0 r_\perp)^2} + \gamma + \log\left(\frac{k_0 r_\perp}{2a}\right)\right]D. \quad (29)$$

The first order potential inside the stream is then given by

$$\Phi_<^{(1)} \simeq \frac{\sigma_0^2}{q_0 a}\left[\log\left(\frac{k_0 r_\perp}{2a}\right) + \gamma\right]D, \quad (30)$$

where we have used $\mu = \pi r_\perp^2\rho$ together with relation (9). The Euler constant is $\gamma = 0.577$.

Using (22) and (30) we obtain a closed set of equations for the perturbations D that can now be linearised. We therefore set $D << 1$, as well as $\langle v^2\rangle(z,t) = \sigma^2(t) + O(v_1^2)$ which gives

$$\ddot{D} + 2\frac{\dot{a}}{a}\dot{D} = \frac{1}{a^2}\partial_z^2\Phi_<^{(1)} - \frac{k_0^2\sigma^2}{a^2}D, \quad (31)$$

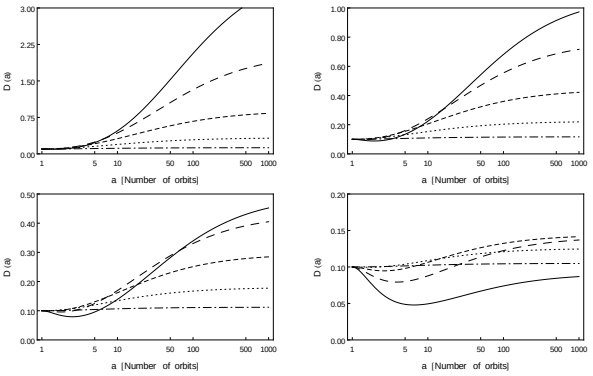

Figure 4. *Dynamics of the perturbation D with respect to the scale factor a for the comoving factors $k_0 r_\perp$ = 0.9 (solid), 0.7 (wide-dashed), 0.5 (narrow-dashed) 0.3 (dotted) and 0.1 (dashed-dotted). From top left to bottom right: q = 0.5, 0.75, 1, 2*

$$\ddot{D} + 2\frac{\dot{a}}{a}\dot{D} = -\frac{k_0^2}{a^2}\left\{\frac{2G\mu_0}{a}\left[\log\left(\frac{k_0 r_\perp}{2a}\right) + \gamma\right] + \sigma^2\right\}D. \quad (32)$$

In a collisionless cylinder the longitudinal velocity dispersion decreases as $\sigma = \sigma_0 a^{-1}$ (see eq. 21), while the perpendicular velocity dispersion stays constant. Equation (32) can therefore be written as

$$\ddot{D} + 2\frac{\dot{a}}{a}\dot{D} = -\frac{\sigma_0^2 k_0^2}{q_0 a^3}\left\{\log\left(\frac{k_0 r_\perp}{2a}\right) + \gamma + \frac{q_0}{a}\right\}D. \quad (33)$$

The perturbation, D, is damped if the right hand side of equation (33) is negative. Therefore we can define a Jeans length

$$\lambda_J = \pi r_\perp \exp\left(\frac{q_0}{a} + \gamma\right), \quad (34)$$

which is very different from the stability criterion in a homogeneous surrounding (14). The geometry of a thin cylinder leads to a Jeans length with an exponential form that guarantees stability up to much larger scales.

Equation (33) can now be simplified using (4) and taking a as variable:

$$D''(a) + \frac{2}{a}D'(a) = \quad (35)$$
$$-\frac{4(k_0 r_\perp)^2}{q_0 a^3}\left\{\log\left(\frac{k_0 r_\perp}{2a}\right) + \gamma + \frac{q_0}{a}\right\}D(a),$$

There are two remaining free parameters, namely q_0 and $k_0 r_\perp$, which describe the eccentricity of the orbit and the scale of the perturbation compared to the width of the stream. In Fig. 4 we plotted the numerical solutions for different sets of parameters. For a small Toomre parameter the perturbations will become nonlinear and we expect gravitational collapse to occur. However, for larger q the perturbations either undergo a damped oscillation or they freeze out after an unsubstantial phase of growth. Comparing these results with the relation (9) leads to the conclusion that the Toomre parameter of a tidal stream is always large enough to assure stability in all cases of interest.

In order to see the effect due to the linear expansion, we also look at the case of a static cylinder. The evolution of perturbations is then given by equation (33) with $a = 1$ and $\dot{a} = 0$ and its behaviour is plotted in Fig. 5. Even for a large Toomre parameter q, there are always collapsing modes supposing an infinitely extended cylinder. This is fundamentally

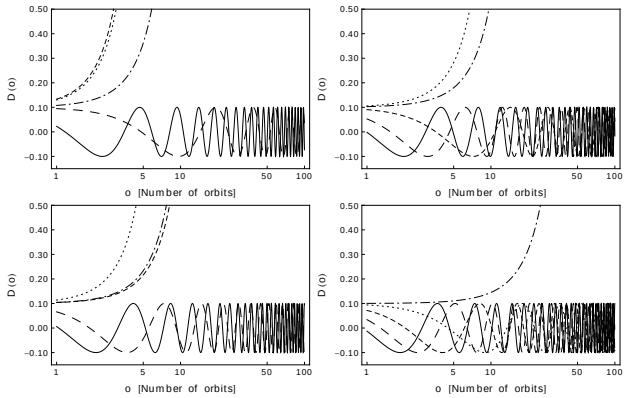

Figure 5. *Perturbations D in the case of a static cylinder, where o is the number of orbits. The different lines represent the factors $kr_\perp = 0.9$ (solid), 0.7 (wide-dashed), 0.5 (narrow-dashed) 0.3 (dotted) and 0.1 (dashed-dotted) in static coordinates. From top left to bottom right: $q = 0.5, 0.75, 1, 2$*

different in the expanding case, where all modes are damped for a high enough q, leading to stability on all scales.

Until now we analysed the stability of a stream with linear perturbation theory. In the next section we take a different look at the stream stability by exploring the longitudinal collapse of cylindrical slices. This somehow more heuristic approach is not restricted to the linear regime and gives an independent analysis of the problem.

4 SHELL COLLAPSE IN A CYLINDER

In a one dimensional case of an extended cylinder the spherical collapse reduces to the longitudinal collapse of thin slices. We therefore consider a homogeneous and infinitely long expanding cylinder with a top hat perturbation at the time t_i. The stream can then be cut into slices, which evolve at constant energy. The energy at a certain distance s is given by

$$E_i = \frac{1}{2}v_i^2 + \Phi(s) = \frac{1}{2}\left(\frac{\dot{a}_i}{a_i}\right)^2 s^2 - \frac{GM_i(s)}{s}. \tag{36}$$

Since the mass M_i evolves as

$$M_i(s) = \int (1+\delta)\mu_b(t_i)ds = \frac{2\mu_0}{a_i}(1+\delta)s \tag{37}$$

we obtain the energy

$$E_i = \frac{1}{2}\alpha^2\left(\frac{s}{a_i}\right)^2 - 2G\frac{\mu_0}{a_i}(1+\delta). \tag{38}$$

Slices with a positive total energy will never collapse and therefore $E_i \geqslant 0$ is our stability condition. Equation (38) then leads to

$$s \geqslant r_\perp\sqrt{\frac{8(1+\delta)a_i}{q_0}}, \tag{39}$$

where we have used the definition of the Toomre parameter (9). Slices further away are stable while nearby ones will collapse. The critical distance below which the stream becomes unstable is growing with the square root of time.

In the picture of shell collapse the velocity dispersion is completely ignored, since the diffusion of particles into other slices makes the problem much more complicated. We will however account for the dispersion by an *ad-hoc* introduction of the Jeans length λ_J, which guaranties the stream stability on small scales. With (39) and (34) we can then construct the stability criterion

$$q_0 \geqslant \frac{8(1+\delta)}{\pi^2}a_i e^{-2(q_0 a_i^{-1}+\gamma)}, \tag{40}$$

which is fulfilled at the beginning ($a = 1$) and may be violated at some later times ($a > a_c$). This means that for $t_i = t_0$ all instable slices are below λ_J and therefore all the stream is stable. Later on however and depending on q_0 unstable modes may appear just above λ_J.

Since the Jeans length gives a minimum size for the final structure, the initial collapse must start at a scale well above this. A calculation of the collapse-time t_{coll} shows however that t_{coll} dramatically grows with the distance of the slice. Slices only a few times further away than the Jeans length already have a t_{coll} that largely exceeds one Hubble time, at least for $q_0 \geqslant 1$. This means that even though there are unstable modes in an expanding stream, they will never have enough time to grow substantially. Collapse only occures for very small values of q_0 well below the limit given by (9).

This qualitative picture is in agreement with the results plotted in Fig. 4, where a phase of damped oscillation is followed by a phase of growth, freezing out at a very low level still in the linear regime.

5 TOWARDS A REALISTIC STREAM

A realistic treatment of a tidal stream orbiting its host galaxy can become very complex, which leads us to consider the possibility that our model of an expanding cylinder is an over-simplification and therefore we are missing some important dynamics. In the following we treat possible deviations to our model and discuss their influence on the stability:

- In general, the host galaxy is not simply isothermal, but can have a triaxial shape that varies with time, and it contains substructures. The orbit of a cluster is then no longer within a plane and it may lose its regularity. The analysis of stability effects in such a complex situation is best tackled with full numerical simulations. Nevertheless, there is no a-prior reason to believe that one of these effects could fundamentally alter the stability criterion.
- A stream approximately traces the orbit of its cluster and is therefore more and more curved the longer it gets. This does not correspond to the straight cylinder used in the model. However the effect of the bending is rather stabilising the stream against longitudinal Jeans instabilities and can therefore confidently be ignored.
- A much more severe limitation of our model is the assumption of a cylindrical form. In reality streams are often more sheet-like and their thickness strongly varies during the orbital period. The closer a stream approaches the centre of the host, the thinner it gets. The reason for this behaviour is the form of the isothermal host potential which leads to orbits that occupy a narrower real-space volume closer to its centre. In Fig. 6 the image of a stream on an eccentric orbit is illustrated. The difference in the stream-width is very pronounced and the sheet like structure at apocenter is also visible. Even though the variation in the thickness has a major influence on the local stream density, it does not affect

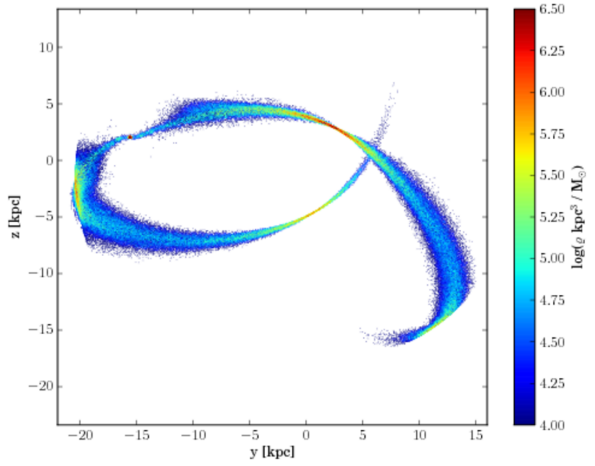

Figure 6. Density map of a star cluster with a leading and tailing stream after 2 Gyr in an isothermal host potential. The orbit lies in the (y,z)-plane and has an eccentricity factor of $b = 0.74$. The high eccentricity leads to strong variations in the stream width. While the streams are narrower and denseer at pericentre, they become flattend at apocentre with the typical umbrella-like form.

the longitudinal collapse condition, which only depends on the linear density. Incorporating the effect of the flattening of the stream is somewhat more difficult because it affects the potential (30). However, it is again unlikely that the sheet-like structure would have an enhancing effect on the collapse since it is stretching the stream which reduces its density.

• As the stream orbits between apocenter and pericenter, its length is oscillating, a fact that is not included in our model assumptions and may affect the stream stability. In fact, the stream only expands linearly on average, its length oscillates during one orbit, being stretched at pericenter and compressed at apocenter. In Fig. 7 the average distance of random points in streams on different orbits are illustrated and the orbital oscillation as well as the overall linear expansion are clearly visible. These oscillations have an effect on the longitudinal perturbations. From peri- to apocenter, when the stream-length is shrinking, we are no longer in a stable regime and we expect growth. However, this growth happens on a timescale longer than the orbital period so that perturbations do not have time to collapse. This can be shown by approximating the shrinking of the stream with a linearly decreasing scale factor of the form

$$a(t) = d - (d-1)\,\frac{2}{T}t, \tag{41}$$

where $d = l_{max}/l_0$. The stream length $r(t) = a(t)l_0$ now runs from l_{max} to l_0 in half of an orbital period. We then use the equation of perturbation (33) and replace the time variable with the scale factor. The result is

$$D''(a) + \frac{2}{a}D'(a) = \tag{42}$$
$$-\frac{(k_0 r_\perp)^2}{(d-1)^2 q_0 a^3}\left\{\log\left(\frac{k_0 r_\perp}{2a}\right) + \gamma + \frac{q_0}{a}\right\} D(a),$$

as well as the initial conditions $D(d) = 0.1$ and $D'(d) = 0$. We find growing solutions if the right hand side of the above

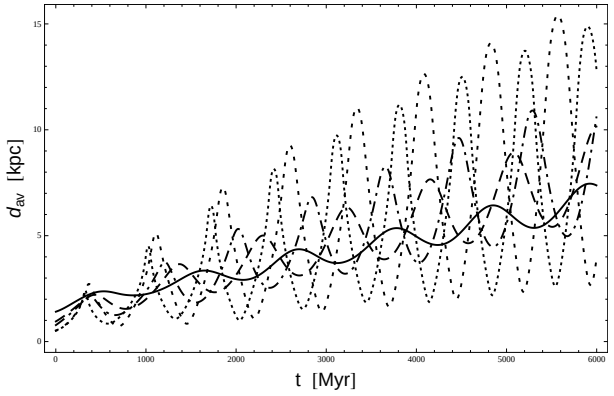

Figure 7. The evolution of the distance between chosen particles in the stream for different simulations with $b = 0.14$ (full), $b = 0.34$ (dashed), $b = 0.54$ (dashed-dotted), $b = 0.74$ (narrow-dotted) and $b = 0.88$ (broad-dotted). Whilst there is linear growth averaged over the orbital motion, the length is oscillating with the orbit, and the amplitude of the oscillation is larger for higher eccentricity.

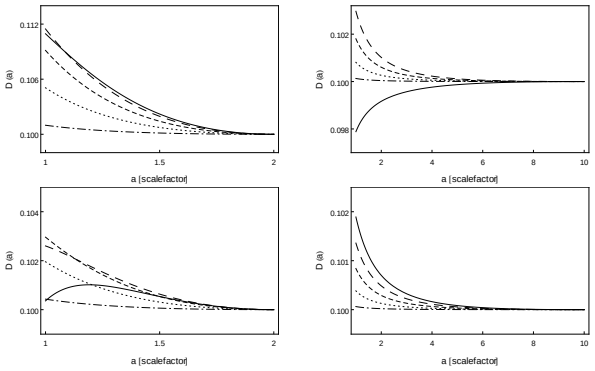

Figure 8. *Growing perturbations D for a shrinking scale factor a with $k_0 r_\perp = 0.9$ (solid), 0.7 (wide-dashed), 0.5 (narrow-dashed), 0.3 (dotted), 0.1 (dashed-dotted). The plots should be read from right to left. Top: $q = 0.5$ with $d = 2$ (left) and $d = 10$ (right). Bottom: $q = 1$ with $d = 2$ (left) and $d = 10$ (right).*

equation is positive, where the actual value determines the growth rate. A large value of d (high eccentricity) gives a small growth factor for a long interval of integration, whilst a small value (low eccentricity) gives a large growth factor for a short interval, the reason being the d^2-term in the denominator of (42). Hence, the actual growth of perturbations stays negligibly small in all cases even for a q as low as 0.5 and the overall stability of our streams is therefore ensured. In Fig. 8 we plotted the evolution of the perturbations between peri- and apocenter for the case of $q = 0.5$ and $q = 1$ and with $d = 2$ and $d = 10$.

• Because of the longitudinal contraction at apocenter and the transversal contraction at pericenter the density and the total velocity dispersion are oscillating twice as fast as the stream length. This can be observed in Fig. 9, where we plotted the longitudinal and transversal velocity dispersion of a stream with high orbital eccentricity. The doubling of the frequency comes from the fact that the stream is longitudinally compressed at apocenter and transversely compressed at pericenter. Fig. 9 can be understood qualitatively by assuming that the particles in the stream are on nearly free

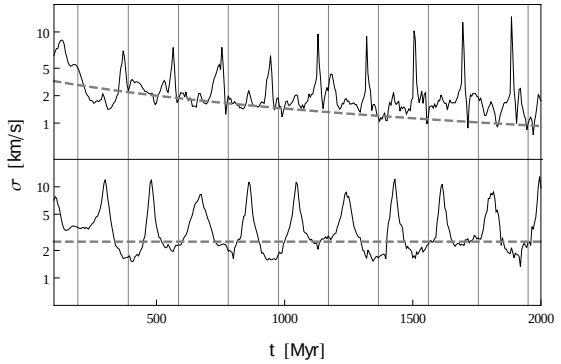

Figure 9. *Evolution of the velocity dispersion for a high eccentricity orbit ($b = 0.75$). The dispersion parallel to the stream is plotted at the top, the one perpendicular to the stream at the bottom. The grey dahsed curves show the time evolution predicted by the model. The vertical lines correspond to the apocenter passage of the cluster.*

epicyclic orbits around the host, which means that the host potential is dominating and that the stream particles are not feeling each other. Slightly displaced orbits are then crossing at apo- and again at pericenter which leads to large peaks in the velocity dispersion.

Our model predicts a longitudinal dispersion that decreases on average, an effect that is not clearly visible in the plot on the top of Fig. 9. Whilst the minima in the longitudinal dispersion seem to decrease as predicted, the maxima are growing with time. This growth comes from the fact that the particle orbits separate more and more to end up at distinct free orbits with the same eccentricity but with a shift in the azimuthal angle. The particles are then all crossing at the same place leading to a sharp peak in the dispersion. The orbital oscillation is also visible in the plot of the transversal dispersion at the bottom of Fig. 9. On average however the transversal dispersion seems to stay constant as predicted by the model. A more detailed study of the stream dispersion was done by Helmi & White (1999), who found a similar evolution of the dispersion over many more orbital periods.

6 CONCLUSIONS

We have studied the gravitational stability of tidal streams by modelling them as thin linearly expanding cylinders of collisionless matter. Such a model leads to a stability criterion that has an exponential dependence on the one dimensional Toomre parameter. We derive a perturbation analysis and also use energetic arguments, to show that a cylinder with the dispersion, the density and the growth rate of a tidal stream is stable for all times.

We used numerical simulations to test our main approximations and to study the detailed phase space evolution of tidal streams. As a final consistency check, we note that none of our simulations show any evidence for gravitational instability.

In reality, a stream is only linearly expanding on average, its length is oscillating during one orbit. This leads to a time interval between apo- and pericenter, where the scale factor shrinks again and the stream is in an unstable regime. Nevertheless, this time interval is too short for the perturbations to grow substantially and the oscillation of the stream length has therefore no influence on the stability.

Collisionless stellar or dark matter streams should therefore evolve smoothly in time, simply stretching further away from the parent system. The structure observed in tidal streams, such as Palomar 5 must have an external origin, perhaps disk shocking or encounters with molecular clouds or dark matter substructures.

Our stability analysis could in principle also be extended to other systems producing streams. However, systems with non spherical shapes and net angular momentum are extremely difficult to analyse with analytical methods, since the alignment of the interacting objects is important. Merging disk galaxies for example produce streams with internal structures strongly depending on the initial alignment of the disks and on their angular momenta. In such cases, high resolution numerical simulations are the indispensable tool for a consistent stability analysis.

ACKNOWLEDGEMENTS

We thank Sebastian Elser and George Lake for helpful discussions. This research is supported by the Swiss National Foundation.

REFERENCES

Barnes, J. E.; Hernquist, L. 1992, Nature, 306, 715
Besla, G., Kallivayalil, N., Hernquist, L., van der Marel, R., Cox, T., Kere, D., 2010, ApJ, 721, L97
Binney, J. & Tremaine, S., 2008, Galactic Dynamics, Princeton Univ. Press, Princeton
Calcneo-Roldn, C., Moore, B., Bland-Hawthorn, J., Malin, D., Sadler, E. M. 2000, MNRAS, 314, 324
Comparetta J. & Quillen, A. C. 2010, preprint(arXiv:astro-ph/1005.4952v1)
Dehnen, W., Odenkirchen, M., Grebel, E. K., & Rix, H.-W. 2004, AJ, 127, 2753
Diemand, J., Kuhlen, M., Madau, P., Zemp, M., Moore, B., Potter, D. & Stadel, J. 2008, Nature, 454, 735
Eyre, A., Binney J. 2011, MNRAS, 413, 1852
Fridman, A. M, & Polyachenko, V. L., 1984, Physics of Gravitating Systems, Volume 1, Springer-Verlag, New York
Grillmair C. J., Freeman, K. C., Irwin, M. & Quinn, P. J. 1995, AJ, 109, 2553
Helmi, A. & White, S. D. M., 1999, MNRAS, 307, 495
Ibata, R. A., Gilmore, G. & Irwin, M. J. 1994, Nature, 370, 6486
Johnston, K. V., Zhao, H., Spergel, D. N., Hernquist, L. 1999, ApJ, 512L, L109
Just, A., Berczik, P., Petrov, M. I. & Ernst, A. 2009, MNRAS, 392, 969
Kallivayalil, N., van der Marel, R. P., Alcock, C. 2006, ApJ, 652, 1213
Klypin, A., Zhao, H., Somerville, R. S. 2002, ApJ, 573, 597
Küpper, A. H. W., MacLeod, A. & Heggie, D. C. 2008, MNRAS, 387, 1248
Law, D. R., Majewski, S. R., Johnston, K. V. 2009, ApJ, 703L, L67

Lin, D. N. C., Lynden-Bell, D. 1977, MNRAS, 181, 59
Lux, H, Read, J. I., Lake, G. 2010, preprint(arXiv:astro-ph/1001.1731v1)
Mayer, L., Moore, B., Quinn, T., Governato, F., & Stadel, J. 2002, MNRAS, 336, 119
Moore, B., Davis, M. 1994, MNRAS, 270, 209
Odenkirchen, M. et al. 2001, ApJ, 548, L165
Odenkirchen, M., Grebel, E. K., Dehnen, W., Rix, H.-W., Cudworth, K. M. 2002, ApJ, 124, 1497
Peebles, P. J. P., 1980, The Large-Scale Structure of the Universe, Princeton Univ. Press, Princeton
Quillen, A. C. & Comparetta J. 2010, preprint(arXiv:astro-ph/1002.4870v1)
Read, J. I., Moore, B. 2005, MNRAS, 361, 971
Schneider, A., Krauss, L., Moore, B. 2010, Phys. Rev. D, 82, 063525
Stadel J. 2001, PhD-thesis, University of Washington
Wetzstein, M., Naab, T. & Burkert, A. 2007, MNRAS, 375, 805
Yoon J. H., Johnston, K. V., Hogg, D. W. 2011, ApJ, 731, Y58
Zemp, M., Moore, B., Stadel, J., Carollo, C. M., Madau, P., 2008, MNRAS, 386, 1543

6

Prospects

There are many further research projects that could extend the work presented in this thesis. This includes studies on the structure formation around the free streaming scale as well as further investigations on collapse within inhomogeneous matter distributions. We will now mention some future projects and discuss the methods to face possible problems.

6.1. Halo mass function around the free streaming scale

In chapter 4 we discussed the intriguing behavior of the halo mass function in ΛWDM cosmologies, and the failure of the EPS approach at scales where free streaming becomes important. An analytical understanding of the mass function around the free streaming scale would however be crucial, because simulations are suffering from severe numerical artifacts.

The most fundamental problem with the EPS approach is the divergence of the predicted mass function at small masses, and this in spite of the fact that matter cannot collapse below the Jeans mass. Including the velocity dispersion in the linear growth calculation as well as modifying the spherical/ellipsoidal collapse (by allowing some kind of shell diffusion) would introduce a cutoff to the EPS mass function and could therefore solve this divergence problem.

The inconsistency of the EPS approach observed in chapter 4 must however have another reason, since the simulations do not include thermal velocities either. Some preliminary calculations also show that the scale where the velocity dispersion should alter the mass function is well below the resolution limit of the simulations. Possible reasons for the discrepancy between the EPS model and the simulations are an inappropriate window function, a modification of the halo ellipticity around the free streaming scale, or a general failure of the EPS approach due to the inversion of the hierarchical structure formation.

In order to find an appropriate description of the mass function around the free streaming scale, more simulations are needed with enough resolution to challenge scales closer to the free streaming limit. It will be interesting to see what kind of qualitative behavior can be observed, namely if the mass function will flatten out further or if it will turn over at some point and and start to decrease.

6.2. Halo concentrations in ΛWDM

The concentration-mass relation in ΛWDM is not a monotonic increasing function as in ΛCDM, but it turns over and decreases again towards small masses. This surprising behavior can be interpreted as an inversion in the hierarchy of structure formation, since the halo concentration is believed to be proportional to the background density at the formation time.

In order to examine the precise hierarchy of halo formation, one would have to analyze the detailed halo accretion history, by tracing back all haloes until the time of virialisation. In this way it would be possible to construct merger trees and to determine the average redshift of halo formation for different halo masses.

An inversion in the formation hierarchy would have an influence on the high redshift evolution of galaxy formation. Since small haloes form later, the beginning of the epoch of reionisation would be delayed, and the first luminous sources would emerge within larger dark matter structures.

A reversed halo formation could also be the reason for the observed discrepancy in the halo mass function, discussed in section 6.1. As a matter of fact the average collapse redshift is predicted by the EPS formalism to depend on the variance of the Gaussian field, and a flat $\sigma(M)$ relation means simultaneous collapse. Since by construction the $\sigma(M)$ relation cannot turn over, an inversion of the hierarchy cannot be described by a simple EPS approach.

6.3. Fragmentation in cosmic filaments

The model of growing perturbations in a cylindrical matter distribution discussed in chapter 5 can also be applied to other systems in astrophysics. One example is the filament as part of the cosmic web, which is spanned between two galaxy clusters and delimitates the large volumes of cosmic voids. Contrary to streams, the filaments are not trapped in large potential wells and do not feel strong tidal forces, leading to further complications like focusing at pericenter and longitudinal oscillations. They are therefore even better suited to the model of cylindrical collapse.

The formation of filaments is usually described by the model of ellipsoidal collapse, where one collapsed axis corresponds to sheet formation, two collapsed axes to filament formation and three collapsed axis to halo formation [43]. The picture of ellipsoidal collapse is however not very consistent with simulations, where structure formation seems to be dominated by voids, and where filaments are stable objects spanned between high density centers. It is therefore worth to analyze whether direct fragmentation along stable filaments could act as an alternative mechanism for halo formation.

The equation of perturbation for an expanding cylinder has been derived in chapter 5. In the case of a general cylindrical potential and a general expansion factor $a(t) \propto t^n$, it has the following form

$$D'' + \frac{(3n-1)}{na} D' = -\frac{4\pi G \Delta \rho_{b0}}{n^2} \left(\frac{a_0}{a}\right)^{\frac{5n-2}{n}} \left\{ \frac{K_1(k_0 r_\perp)}{W(k_0 r_\perp)} + \frac{(k_0 r_\perp)^2 \sigma_0^2}{4\pi a G \Delta \rho_{b0}} - 1 \right\} D, \quad (6.1)$$

where Δ is the virial overdensity of the filament and σ_0 is its longitudinal velocity dispersion. The functions K_1 and W as well as $r_\perp$ and k_0 are definied in chapter 5.

In order to solve Eq. (6.1), one needs to know the longitudinal expansion, the initial density field, and the velocity dispersion of the filament. Following the picture of an ideal ellipsoidal collapse, where different axes collapse at different timescales, a filament corresponds to a structure with two collapsed axes out of three. The expansion factor $a(t)$ should therefore be close to the expansion of the universe, since the third axis is expanding with the cosmic flow. The velocity dispersion on the other hand should correspond to the primordial one, because there is no relaxation or two body interaction in a collisionless fluid and the virial motion should not affect the longitudinal direction. For a ΛCDM cosmology this would mean that σ_0 is completely negligible. The overdensity field should finally correspond to the primordial one, since an ideal collapse of two axes is increasing the density field, leaving the overdensities unchanged.

These approximations are valid for an idealised ellipsoidal collapse, but things can be very different in a real cosmological situation. Deviations from the form of a perfect ellipsoid and tidal forces from the surroundings could for example significantly boost the velocity dispersion and wash out the primordial overdensities. A detailed study of filaments and their virialisation is necessary to determine the importance of these effects.

Proceeding with the above approximations, Eq. (6.1) can be simplified to

$$D'' + \frac{3}{2a} D' = -9\pi G \Delta \rho_{b0} \left(\frac{a_0}{a}\right)^2 \left\{ \frac{K_1(k_0 r_\perp)}{W(k_0 r_\perp)} - 1 \right\} D, \quad (6.2)$$

where we have assumed $n = 2/3$ and where the initial conditions are given by $D_i = \delta(z_c, k)$ with the collapse redshift z_c. This equation has growing modes on all scales and halo formation could be possible, at least if there is enough time for the collapse.

Bibliography

[1] Komatsu, E., Smith, K. M., Dunkley, J., & The WMAP Team. 2011, Ap. J. S, 192, 18

[2] Tegmark, M., & The SDSS Team. 2006, Phys. Rev. D, 74, 123507

[3] Busha, M. T., Adams, F. C., Wechsler, R. H., & Evrard, A. E., 2003, Ap. J. 596 713

[4] Ostriker, J. P., Peebles, P. J. E. & Yahil, A., 1974, Ap. J. 193 L1

[5] Einasto, J., Kaasik, A., & Saar, N., 1974, Nature 250 309

[6] Roberts, M. S. & Rots, A. H., 1973, A. & A. 26 483

[7] White, S. D. M. & Rees, M. J., 1978, MNRAS 183 341

[8] Gehrstein, S. S. & Zeldovich, Y. B., 1966, JETPL 4 120G

[9] Reines, F., Sobel, H. W., & Pasierb, E., 1980, Phys. Rev. Lett. 45 1307

[10] Peebles, P., 1982, Ap. J. Lett. 263 L1

[11] D'Amico, G., Kamionkowski, M. & Sigurdson, K., 2009, arXiv:0907.1912

[12] Roos, M., 2010, arXiv:1001.0316

[13] Milgrom, M., 1983, Ap. J. 270 365M

[14] The CDMS Collaboration: Ahmed, Z. *et. al.*, 2009, Phys. Rev. Lett. 106 131302

[15] The EDELWEISS Collaboration: Armengaud, E. *et. al.*, 2011, Phys. Lett. 702 329

[16] The XENON Collaboration: Aprile, E. *et. al.*, 2011, Phys. Rev. D 84 061101

[17] Bertone, G. & Merritt, D., 2005, Mod. Phys. Lett. A, 20 1021B

[18] Abazajian, K. N., *et al.*, 2012, arXiv:1204.5379

[19] Dodelson, S. & Widrow, L. M., 1994, Phys. Rev. Lett. 72 17

[20] Seljak, U., Makarov, A., McDonald, P., & Hy, T., 2006, Phys. Rev Lett. 97 191303

[21] Viel, M., Lesgourgues, J., Haehnelt, M. G., Matarrese, S., & Riotto, A., 2006, Phys. Rev. Lett. 97 071301

[22] Boyarsky, A., Lesgourgues, J., Ruchayskiy, O., Viel, M., 2009, Phys. Rev. Lett. 102 201304

[23] Ellis, J., Hagelin, J. S., Nanopoulos, D. V., Olive, K., & Srednicki, M. 1984, Nuclear Physics B, 238, 453

[24] Mo, H., van den Bosch, F. C., & White, S. D. M., 2010, Cambridge University Press

[25] Peebles, P. J. P., 1980, Princeton: University Press

[26] Padmanabhan, T., 1995, Cambridge University Press

[27] Carroll, S. M., 2004, Addison Wesley, San Francisco

[28] Mukhanov V.F., Chibisov G.V., 1981, Soviet Jour. of Exper. and Theor. Phys. Lett., 33, 532

[29] Bardeen, J. M., Bond, J. R., Kaiser, N., Szalay, A. S., 1986, ApJ, 304 15

[30] Viel, M., Lesgourgues, J., Haehnelt, M. G., Matarrese, S., & Riotto, A. 2005, Phy. Rev. D, 71, 063534

[31] Barnes, J., Hut, H., 1986, Nature, 324 446B

[32] Stadel, J. G. 2001, PhD thesis, University of Washington

[33] Press, W. H., Schechter, P., 1974, ApJ, 187 425P

[34] Zentner, A. R. 2006, IJMPD, 16 763Z

[35] Sheth, R. K., Tormen, G. 1999, MNRAS, 308 119S

[36] Loeb, A. & Zaldarriaga, M., 2005, Phys. Rev. D 71 103520

[37] Bertschinger, E., 2006, Phys. Rev. D, 74 063509

[38] Diemand J., Moore B., & Stadel J., 2006, Nature, 433 389

[39] Schneider, A., Krauss L. M., & Moore, B., 2010, Phys. Rev. D 82 063525

[40] Viel, M., Markovic, K., Baldi, M., & Weller, J. 2011, MNRAS, 421 50V

[41] Schneider, A., Smith, R. E., Macció, A. V., & Moore, B. 2012, MNRAS, tmp 3197S

[42] Schneider, A. & Moore, B. 2011, MNRAS, 415, 1569S.

[43] Shen, J., Abel, T., Mo, H. J., & Sheth, R. K., 2006, ApJ, 645 783S